

STIC-Biotech/ChemLib

105628

Fr m: Hines, Ja-Na
Sent: Wednesday, October 08, 2003 1:29 PM
T : STIC-Biotech/ChemLib
Subject: Sequence search

RECEIVED

OCT -8 2003

Good Afternoon,
Could you please search amino acids 1 to 63 of SEQ ID NO:1, SEQ ID NO:1,
1 to 63 of SEQ ID NO:2 and 64-456 of SEQ ID NO:2;
1 to 672 of SEQ ID NO:3 and SEQ ID NO:3, and
1 to 672 and 693 to 1086 of SEQ ID NO:4 all for 09/147,052.
Thanks!!
Ja-Na Hines (76048)
CM1 9A11
AU:1645

STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 10/8/03
Date Completed: 10/10/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: *Reverse to AA*
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: *OX*
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 105628

TO: Jana Hines
Location: CM1/9A11/7E12
Art Unit: 1645
Friday, October 10, 2003

Case Serial Number: 09/147052

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Hines,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

- TD
- W3

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:02 ; Search time 58.5355 Seconds
(without alignments)
5924.998 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200

Sequence: 1 atgactattttaggcgaa.....catattttgcaatggcccaat 672

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-rlp
-Q/cgn2.1/USPTO.spool/US09147052/runat_08102003_154341_29846/app_query.fasta_1.5980
-DB-SPREMBL_23 -QFMT-fastan -SUFFIX-rspt -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT-pco -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09147052 @CGN_1_1_771 @runat_08102003_154341_29846 -NCPU=6 -ICPD=3
-NO_MMAL -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1200	100.0	805	12	Q98Y44 turkey herp

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45			
Q98Y45	100.0	100.0	100.0	100.0	99.4	98.7	84.3	84.3	83.8	82.5	85.2	88.1	87.9	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	504	501	501	500	500	500	500	500	500	500	500	500			
Q8JLW3	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500		
Q8JLW4	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	
Q8JLW2	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614	51.2	607	605	590	588.5	583	573	568	563.5	559.5	526	516	513	505	505	504	504	501	501	500	500	500	500	500	500	500	500	500	500	500	500	500
Q83291	805	865	865	865	865	865	870	870	870	865	848	881	879	868	614</																																

Pred. No.:	3.03e-118	Length:	805
Score:	1200.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-147-052-3 COPY 1.672 (1-672) x Q98Y44 (1-805)

Qy	1	ATGCACTATT	TTAGCGGA	TTCGATAT	TTTCCCTAT	TATAGTAT	TCTATAT	CGGTACG	AAC	60
Db	1	MethIstYrPheArgAsnCysIlePhePheLeuIleValIleLeuYrGlyThrAsn	20							
Qy	61	TCATCTCGGAGTACC	AAAAATGTCACATCAAGAGAAGTGT	TTTCGAGCGTC	CACGGTGTCT	120				
Db	21	SerSerProSerThrGlnAsnValThrSerArgGluValSerSerValGlnLeuSer	40							
Qy	121	GAGGAAGAGTCTACG	TTTTTATCTTTGTGCCCCCACGAGTG	GGTTCACCGTGAT	CCGCTCTA	180				
Db	41	GluGluGluSerThrPheTyrlEuCysProProValGlySerThrValIleArgLeu	60							
Qy	181	GAACCGCGCGAAATATG	CCCCGAACCTAGAAAACCCACCGAGTGGGTGAAGGAATCGCG	240						
Db	61	GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla	80							
Qy	241	ATATATTATAAGAGAATAT	CAGTCCCATATAAATTTAAAGTAGCGCTTTATATAAAAT	300						
Db	81	IleLeuPheLysGluAsnIleSerProTyrlLysPheLysValThrLeutyrylLysAsn	100							
Qy	301	ATCATTACAGCACGACAT	TGGACGGGACGACATATAGACAGATCACTAATCGATATACA	360						
Db	101	IleIleGluThrThrThriPhgGlyThrThryrargGlnIleThrAsnArgTyThr	120							
Qy	361	GATPAGCACCGCGTTTCC	ATTGAAGAGATCACGGATCTAATPCGACGGCGCAAAGGAATGC	420						
Db	121	AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys	140							
Qy	421	TCATCTTAAGCAGATACCT	TAGAACCAATCTATATGTTGAAGCGTTTGACAGGGATGCG	480						
Db	141	SerSerLysAlaArgtyrLeuargAsnAsnValTyrralGluAlaPheAspArgaspAla	160							
Qy	481	GGAGAAAAACAAGTACTT	CTTAACACCATCAAAATTCACACGCGCCCGAATCTAGGGCATGG	540						
Db	161	GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180							
Qy	541	CACACGACTAATGAGACGTAT	ACCCTGTGGGATCACCATCGGATATATCCAACGGGAACC	600						
Db	181	HisThrThrAsnGluThrTyrrThrValTrpGlySerProTripletTyArgthrGlyThr	200							
Qy	601	TCGCTCAANTGTATAGTAG	AGAAATGGATGCGCGCTCTGTGTTTCGGTATTCATATTTT	660						
Db	201	SerValAsnCysIleValGluMetAspAlaArgSerValPheProTyrlSerTyrlPhe	220							
Qy	661	GCAATGGCCNAAT	672							
Db	221	AlaMetaIalaasn	224							

RESULT 2

Q89Y41	Q89Y45	PRELIMINARY;	PRT;	805 AA.
Q89Y45	Q89Y45;			
AC	Q89Y45;			
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)		
DE	Glycoprotein B (virus)			
OS	Turkey herpesvirus.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae.			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=10390;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RA	Davidson I.;			
RT	"Marek's disease virus - chicken isolate, gB gene."			

DE Glycoprotein B.
 GN GB.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-N;
 RA Parcellis M.S., Shamblin C.E., Dienglewicz R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 RT Pathotypes of Marek's Disease Viruses (vMDV, vVMDV, vv+MDV): Mutations
 RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129967; AA097700.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98030 MW; 9008B58B4B4E114D CRC64;

Alignment Scores:
 Pred. No.: 3.08e-118 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x Q8JLW3 (1-865)

QY 1 ATGCCTATTTAGCGGAATTCATATTTTCCCTATAGTTATCTATATGTTACGACGAC 60
 Db 1 MetHisTyrPheArgAsnGlnCysIlePhePheLeuValIleLeuTyrGlyThrAsn 20

QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTGTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40

QY 121 GAGGAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCACCGTGTATGCTCTA 180
 Db 41 GluGluGluSerThrPheTyrLeuGlnCysProProValGlySerThrValIleArgLeu 60

QY 181 GAACGCCCGCGAAATGTCCTCAACCTAGAAAGCCAGCGTGGGTGAAGAAATCGCG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80

QY 241 ATATTATTTAAAGAGATATCATGTCATATAATTTAAAGTCAAGCTTTTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100

QY 301 ATCATTCAGACGACATGACGGGAGCAGATATAGACATCACTAATCGATATACA 360
 Db 101 IleIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120

QY 361 GATAGGACCCCGTTTCCATTTGAGAGATCAAGAGATCAAGAGATCAAGAGATCAAG 420
 Db 121 AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys 140

QY 421 TCATCTAAAGCAAGATACCTTTAGAACATGATGATGATGATGATGATGATGATGAT 480
 Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValIleAlaPheAspArgAspAla 160

QY 481 GGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGCGAATCTAGGCGATGG 540
 Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180

QY 541 CACAGCACTAATGACAGCTATACCGTGTGGGATCACCATGATGATATATCGACGGAGCC 600
 Db 181 HisThrThrAsnGluThrTyrThrValTyrGlySerProTyrPheArgThrGlyThr 200

QY 601 TCCGTCAATGTATAGTACGAAATGGATGCCGCTCTGCTTTCCGCTATTCATATTTT 660
 Db 201 ServaAsnGlnValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220

QY 661 GCAATGGCCAAT 672
 Db 221 AlaMetAlaAsn 224

RESULT 4
 Q8JLW4
 ID Q8JLW4 PRELIMINARY; PRT; 865 AA.
 AC Q8JLW4;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X;
 RA Parcellis M.S., Shamblin C.E., Dienglewicz R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 RT Pathotypes of Marek's Disease Viruses (vMDV, vVMDV, vv+MDV): Mutations
 RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129965; AA097698.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98107 MW; 4D2628B5E4DEB28B CRC64;

Alignment Scores:
 Pred. No.: 3.08e-118 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x Q8JLW4 (1-865)

QY 1 ATGCCTATTTAGCGGAATTCATATTTTCCCTATAGTTATCTATATGTTACGACGAC 60
 Db 1 MetHisTyrPheArgAsnGlnCysIlePhePheLeuValIleLeuTyrGlyThrAsn 20

QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTGTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40

QY 121 GAGGAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCACCGTGTATGCTCTA 180
 Db 41 GluGluGluSerThrPheTyrLeuGlnCysProProValGlySerThrValIleArgLeu 60

QY 181 GAACGCCCGCGAAATGTCCTCAACCTAGAAAGCCAGCGTGGGTGAAGAAATCGCG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80

QY 241 ATATTATTTAAAGAGATATCATGTCATATAATTTAAAGTCAAGCTTTTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100

QY 301 ATCATTCAGACGACATGACGGGAGCAGATATAGACATCACTAATCGATATACA 360
 Db 101 IleIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120

QY 361 GATAGGACCCCGTTTCCATTTGAGAGATCAAGAGATCAAGAGATCAAGAGATCAAG 420
 Db 121 AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys 140

QY 421 TCATCTAAAGCAAGATACCTTTAGAACATGATGATGATGATGATGATGATGATGAT 480
 Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValIleAlaPheAspArgAspAla 160

```

QY 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATCAACAGCCGCCGAATCTAGGCGATGG 540
D 161 GlycyluysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTyr 180
QY 541 CACAGGACTAATGAGAGATACCGTGTGGGATCACCAGTGGATATATCGAAGCGGAACC 600
D 181 HisthrThrAsnGluThrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCCGTCGAATGTATAGTAGAGAAATGATGCCCGCTCTGTGTTCGCTATTTCATATTTT 660
D 201 SerValasnCysIleValGluGluMetaspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGGCCCAAT 672
D 221 AlaMetaAsn 224

RESULT 5
O8JLW2 PRELIMINARY; PRT; 865 AA.
AC O8JLW2;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=attenuated GA;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (vMDV, vVMDV, vv+MDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129989; AAM97702.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98020 MW; F4F526114EA02DE5 CRC64;

```

```

Alignment Scores:
Pred. No.: 1.7e-117 Length: 865
Score: 1193.00 Matches: 223
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 99.42% Indels: 0
DB: 12 Gaps: 0

```

US-09-147-052-3_COPY_1_672 (1-672) x O8JLW2 (1-865)

```

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTCTATATGTCAGCAAC 60
D 1 MethistyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTGTCGAGCGTCAGTGTCT 120
D 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGCAAGACGTACGTTTATCTTGTGTCCTCCACAGTGGTTCACCCGTCATCCGTC 180
D 41 GluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAACCGCGCGGAAAATGTCCCGAACCTAGAAAACCCAGTGGGTGAGGAATCGCG 240
D 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
QY 241 ATATTATTTAAAGAGATATCAGTCCCATATAAAATTTAAAGTCAGCGCTTTATTA 300
D 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100

```

```

QY 301 ATCATTTCAGACGACATGGACGGGACACATATAGACAGATCACTAATCGATATACA 360
D 101 IleIleGlnThrThrThrTrpThrGlyThrThrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGGACGCCGTTTCCATTGAAGAGATCAGCGATCTAATCGACGGCAAGAGATGC 420
D 121 AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
QY 421 TCATCTAAACAAGATACCTTAGAAACAATGATATCTTGAAGCGTTTGACAGCGATCG 480
D 141 SerSerLysAlaArgTyrLeuArgAsnValIleValGluAlaPheAspArgAspAla 160
QY 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATCAACAGCCGCCGAATCTAGGCGATGG 540
D 161 GlycyluysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTyr 180
QY 541 CACAGGACTAATGAGAGATACCGTGTGGGATCACCAGTGGATATATCGAAGCGGAACC 600
D 181 HisthrThrAsnGlyThrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCCGTCGAATGTATAGTAGAGAAATGATGCCCGCTCTGTGTTCGCTATTTCATATTTT 660
D 201 SerValasnCysIleValGluGluMetaspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGGCCCAAT 672
D 221 AlaMetaAsn 224

RESULT 6
O83291 PRELIMINARY; PRT; 865 AA.
AC O83291;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Glycoprotein 100 precursor.
GN GB.
OS Marek disease virus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=38013;
RN [1]
RP SEQUENCE OF 1-498 FROM N.A.
RC STRAIN=JM;
RA Souloparov M.A., Bakhtina M.M., Krendelshtchikov A.V., Babkin I.V.;
RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
RT complex gp100, gp60, gp48 (b-antigen)";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X91985; CAA63039.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;

Alignment Scores:
Pred. No.: 1.54e-116 Length: 865
Score: 1184.00 Matches: 222
Percent Similarity: 99.11% Conservative: 0
Best Local Similarity: 99.11% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x O83291 (1-865)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTCTATATGTCAGCAAC 60
D 1 MethistyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTGTCGAGCGTCAGTGTCT 120

```



```

Db      21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY      121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCACACCGTGCATCGTCTA 180
Db      41 GluGluGluSerThrPheTyrLeuGlySerProProValGlySerThrValIleArgLeu 60
QY      181 GNACCGCGCGGAAATGTCCCGAACCTAGAAAGCCAGCCAGTGGGGTGAGGAATCGCG 240
Db      61 GluProProArgLysCysLeuGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
QY      241 ATATTATTAAAGACAAATCAGTCCATATAAATTTAAAGTGCACGCTTTATTATAAAAT 300
Db      81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY      301 ATCATTCAGACGACGATGACGGGGAGCAGCATATAGACAGATCACTAATCGATATACA 360
Db      101 IleIleGlnThrThrTriThrGlyThrIleTyrArgGlnIleThrAsnArgTyrThr 120
QY      361 GATAGAGCGCGTTTCCATTGAAGAGATCAGGATCTAATCCAGCGCAAGAGAGATGC 420
Db      121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
QY      421 TCATCTAAAGCAAGATACCTTAAACCAATCAAAATTCACACGCCCGCAATCTAGGCGATGG 480
Db      141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAlaTrp 160
QY      481 GGAGAAAACAAGTACTTCTAAACCAATCAAAATTCACACGCCCGCAATCTAGGCGATGG 540
Db      161 GlyGluLysGlnValLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY      541 CACAGCACTAATGACAGCATACCGTGTGGGATCACCATGGATATATCGAAGCGGAACC 600
Db      181 HistThrThrAsnGluThrTyrThrValTrpGlySerProTyrPheValPheProTyrSerTyrPhe 660
QY      601 TCCGTCATTTGATAGTGTAGAGAAATGGATGCCCTCTGTCTTCCGTTATTTCATATTTT 660
Db      201 SerValAsnCysIleValGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY      661 GCAATGGCCAAT 672
Db      221 AlaMetAlaAsn 224

```

RESULT 7

```

Q69408 ID Q69408 PRELIMINARY; PRT; 864 AA.
AC Q69408;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GB homolog.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RX MEDLINE=9423711; PubMed=8178437;
RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
RT identification and expression by recombinant fowlpox viruses.";
RL Virology 200;484-493(1994).
DR EMBL; U01887; AAA19447.1; -.
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein.B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.
SQ SEQUENCE 864 AA; 98188 MW; 74060C71948F4B7B CRC64;

```

Alignment Scores:

```

Pred. No.: 2,71e-98 Length: 864
Score: 1012.00 Matches: 183
Percent Similarity: 90.18% Conservative: 19
Best Local Similarity: 81.70% Mismatches: 22

```

```

Query Match: 84.33% Indels: 0
DB: 12 Gaps: 0
US-09-147-052-3_COPY_1_672 (1-672) x Q69408 (1-864)
QY 1 ATGCACACTATTAGCGCGGAATTCATATATTTTCCCTATATAGTATTCTATATAGGTACGAAAC 60
Db 1 MetLysTyrPheAsnArgSerLeuPheIlePheLeuThrProIleLeuSerIleAlaThr 20
QY 61 TCATCTCCGAGTACCCAAAATGTGACATCAACAGAGAGTGTTCAGCGGTCCAGTGTGCT 120
Db 21 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCACACCGTGCATCGTCTA 180
Db 41 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 60
QY 181 GNACCGCGCGGAAATGTCCCGAACCTAGAAAGCCAGCCAGTGGGGTGAGGAATCGCG 240
Db 61 GluProProArgLysCysProGluProLeuLysSerThrGluTrpGlyGluGlyIleAla 80
QY 241 ATATTATTAAAGACAAATCAGTCCATATAAATTTAAAGTGCACGCTTTATTATAAAAT 300
Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCAGACGACGATGACGGGGAGCAGCATATAGACAGATCACTAATCGATATACA 360
Db 101 ValIleGlnThrThrTriThrGlyThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGAGCGCGTTTCCATTGAAGAGATCAGGATCTAATCCAGCGCAAGAGAGATGC 420
Db 121 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysGlyLysCys 140
QY 421 TCATCTAAAGCAAGATACCTTAAACCAATGTATATGTTGAAGCGTTTCACAGGATGG 480
Db 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValAspAlaTyrAspArgAspGlu 160
QY 481 GGAGAAAACAAGTACTTCTAAACCAATCAAAATTCACACGCCCGCAATCTAGGCGATGG 540
Db 161 AsnGluLysGlnValLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTrp 180
QY 541 CACAGCACTAATGACAGCATACCGTGTGGGATCACCATGGATATATCGAAGCGGAACC 600
Db 181 HistThrThrAsnGluThrTyrThrValTrpGlySerProTyrPheValTrpArgThrGlyThr 200
QY 601 TCCGTCATTTGATAGTGTAGAGAAATGGATGCCCTCTGTCTTCCGTTATTTCATATTTT 660
Db 201 SerValAsnCysIleValGluMetAspAlaArgSerAlaPheProTyrThrTyrPhe 220
QY 661 GCAATGGCCAAT 672
Db 221 AlaMetAlaAsn 224

```

RESULT 8

```

Q69408 ID Q69408 PRELIMINARY; PRT; 870 AA.
AC Q69408;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UL27 virion membrane glycoprotein B.
GN HVT035.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RX MEDLINE=20578232; PubMed=11134310;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
RT "The genome of turkey herpesvirus.";
RL J. Virol. 75:971-978(2001).

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF291866; AAG45765.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 870 AA; 98816 MW; F93DIC7036938FB1 CRC64;

Alignment Scores:
Pred. No.: 2,71e-98 Length: 870
Score: 1012.00 Matches: 183
Percent Similarity: 90.18% Conservative: 19
Best Local Similarity: 81.70% Mismatches: 22
Query Match: 84.33% Indels: 0
DB: 12

US-09-147-052-3_COPY_1_672 (1-672) x Q9DPO9 (1-870)
QY 1 ATGCACATATTTAGCGGAATTCATATTTCTTATAGTATTTCTATATGTTAGGACGAC 60
DB 7 MetLysTyrPheAsnArgSerLeuPheIlePheLeuThrProIleLeuSerIleAlaThr 26
QY 61 TCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCAGTGTGCT 120
DB 27 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 46
QY 121 GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCACCGTATCCGCTCTA 180
DB 47 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 66
QY 181 GAACCGCGCGAATATCCCGAACCTAGAACCCAGCCAGTGGGTGAAGGAATCCGC 240
DB 67 GluProArgLysCysProGluProLeuLysSerThrGluThrValThrAlaArgGluIleValSer 86
QY 241 ATATTATTTAAAGAGAATATCCAGTCCCATATAAAATTTAAAGTGGCTTTATATAAAAT 300
DB 87 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 106
QY 301 ATCATTACAGCAGCAGATGACGGGACGACATATAGACAGATCACTAATCGATATACA 360
DB 107 ValIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 126
QY 361 GATAGGACGCGCTTCCATTGAAGAGATCAGGATCTAATCGACGCGGAAAGGATGC 420
DB 127 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysCys 146
QY 421 TCATCTAAAGCAAGATACCTTAGAACAAATGTTATGTTGAAGCGTTTACAGGATCGC 480
DB 147 SerSerLysAlaArgTyrLeuArgAsnValThrValAspAlaTyrAspArgAspGlu 166
QY 481 GGAAAAACAAGTACTTCTAAACCATCAAAATTTCAACGCCCGCCGATCTAGGCGATGG 540
DB 167 AsnGluLysGlnValLeuLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTrp 186
QY 541 CACACGACTAATGAGAGATACCGTGTGGGATCACCATGGATATATCGAACCGGAACC 600
DB 187 HisThrThrAsnGluThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 206
QY 601 TCCGCTCAATTGTATAGTAGGAATAGGATCCCGCTCTGTTTCCGTTATTCATATTTT 660
DB 207 SerValAsnCysIleValGluGluMetAspAlaArgSerAlaPheProTyrThrTyrPhe 226
QY 661 GCATGCGCCAAAT 672
DB 227 AlaMetAlaAsn 230

RESULT 9
Q9E1G4 PRELIMINARY; PRT; 870 AA.
ID Q9E1G4
AC Q9E1G4;

```

```

DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE UL27 glycoprotein gb.
GN UL27.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RX MEDLINE-21195611; PubMed-11297687;
RA Kingham B.F., Zelnik V., Kopacek J., Majerciak V., Ney E.,
RA Schmidt C.J.;
RT "The genome of herpesvirus of turkeys: comparative analysis with
RT Marek's disease viruses.";
RL J. Gen. Virol. 82:1123-1135(2001).
DR EMBL; AF282130; AAG30067.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 870 AA; 98827 MW; CE02FB7E7F9EA4A2 CRC64;

Alignment Scores:
Pred. No.: 2,71e-98 Length: 870
Score: 1012.00 Matches: 183
Percent Similarity: 90.18% Conservative: 19
Best Local Similarity: 81.70% Mismatches: 22
Query Match: 84.33% Indels: 0
DB: 12

US-09-147-052-3_COPY_1_672 (1-672) x Q9E1G4 (1-870)
QY 1 ATGCACATATTTAGCGGAATTCATATTTCTTATAGTATTTCTATATGTTAGGACGAC 60
DB 7 MetLysTyrPheAsnArgSerLeuPheIlePheLeuThrProIleLeuSerIleAlaThr 26
QY 61 TCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCAGTGTGCT 120
DB 27 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 46
QY 121 GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCACCGTATCCGCTCTA 180
DB 47 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 66
QY 181 GAACCGCGCGAATATCCCGAACCTAGAACCCAGCCAGTGGGTGAAGGAATCCGC 240
DB 67 GluProArgLysCysProGluProLeuLysSerThrGluThrValThrAlaArgGluIleValSer 86
QY 241 ATATTATTTAAAGAGAATATCCAGTCCCATATAAAATTTAAAGTGGCTTTATATAAAAT 300
DB 87 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 106
QY 301 ATCATTACAGCAGCAGATGACGGGACGACATATAGACAGATCACTAATCGATATACA 360
DB 107 ValIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 126
QY 361 GATAGGACGCGCTTCCATTGAAGAGATCAGGATCTAATCGACGCGGAAAGGATGC 420
DB 127 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysCys 146
QY 421 TCATCTAAAGCAAGATACCTTAGAACAAATGTTATGTTGAAGCGTTTACAGGATCGC 480
DB 147 SerSerLysAlaArgTyrLeuArgAsnValThrValAspAlaTyrAspArgAspGlu 166
QY 481 GGAAAAACAAGTACTTCTAAACCATCAAAATTTCAACGCCCGCCGATCTAGGCGATGG 540
DB 167 AsnGluLysGlnValLeuLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTrp 186
QY 541 CACACGACTAATGAGAGATACCGTGTGGGATCACCATGGATATATCGAACCGGAACC 600
DB 187 HisThrThrAsnGluThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 206

```

QY 601 TCCGTCAATTCATAGTAGAGGAATGGATGCCCGCTCTGTGTTCCGTATTCATATATTT 660
 Db 207 SerValAsnCysIleValGluGluMetAspAlaArgSerAlaPheProTyrThrTyrPhe 226
 QY 661 GCAATGGCCAAT 672
 Db 227 AlaMetAlaAsn 230

RESULT 10

ID Q9PWZ1 PRELIMINARY; PRT; 865 AA.
 AC Q9PWZ1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homolog of HSV-1 major DNA-binding protein (Glycoprotein B).
 GN ORF 35 OR UL27.
 OS Turkey herpesvirus,
 OS Marek's disease virus serotype 2 MDV2, and
 OS Gallid herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390, 36353, 35250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Turkey herpesvirus; STRAIN=HPRS24;
 RA Kato K., Jang H., Izumiya Y., Cai J., Tsuchiwa Y., Miyazawa T.,
 RA Kai C., Mikami T.;
 RT "Identification and Transcriptional Analysis of the Marek's Disease
 RT Virus Serotype 2 Genes Homologous to the Glycoprotein B (UL27), the
 RT ICp18.5 (UL28) and the Major DNA-binding Protein (UL29) Genes of
 RT Herpes Simplex Virus Type 1.";
 RL J. Vet. Med. Sci. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 1; STRAIN=HPRS24;
 RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
 RA Lee Y., Kai C., Takahashi E., Mikami T.;
 RT "The complete DNA sequence and transcription map of the unique long
 RT genome region of Marek's disease virus type 2.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 3; STRAIN=HPRS24;
 RA Izumiya Y., Jang H., Ono M., Mikami T.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 3; STRAIN=HPRS24;
 RA Izumiya Y., Jang H., Ono M., Mikami T.;
 RT "A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
 RT Strain HPRS24.";
 RL Curr. Top. Microbiol. Immunol. 0:0-0(2000).
 DR EMBL; AB024711; BAA83751.1; -;
 DR EMBL; AB024414; BAA82923.1; -;
 DR EMBL; AB049735; BAB16537.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW DNA-binding.
 SQ SEQUENCE 865 AA; 97309 MW; A20E66C70AA75C35 CRC64;

Alignment Scores:

Pred. No.: 1,32e-97 Length: 865
 Score: 1005.50 Matches: 184
 Percent Similarity: 90.67% Conservative: 20
 Best Local Similarity: 81.78% Mismatches: 20
 Query Match: 83.79% Indels: 1
 DB: 12 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x Q9PWZ1 (1-865)

QY 1 ATGCACATATTTAGCGGAATTGCATA---TTTTTCCTTATAGTATTATCTATATGCGTACG 57
 Db 1 MetAsnHisPheSerGlyIleCysValProLeuPheIleSerValValPheTyrPhePhe 20
 QY 58 AACTCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTG 117
 Db 21 GlyArgValAlaArgAlaGlnAsnValThrSerArgGluAlaValSerSerValGlnLeu 40
 QY 118 TCTGAGGAAGAGTCTAGCTTTTATCTTTGTGCCCCCAGTGGTTCACACCGTGATCCGT 177
 Db 41 SerGluAspGluSerSerPhePheLeuCysProAlaValGlyThrValValArg 60
 QY 178 CTAGAACCGCGCGGAAAATGTCCCGAACCTAGAAAAGCCAGTGGGTGAAGGAATC 237
 Db 61 LeuGluProProArgLysCysProGluProLeuLysAlaThrGluTyrGlyGluGlyIle 80
 QY 238 GCGATATATTAAAGAGAATATCAGTCCCATATAAATTAAAGTGACGCTTTATTATAAA 297
 Db 81 AlaIleLeuPheLysGluAsnIleAsnProTyrLysPheLysValThrLeuTyrTyrLys 100
 QY 298 AATATCATTCAGACGACGACATGACGGGGAGCAGACATATAGACAGATCACTAATCGATAT 357
 Db 101 AsnValIleGlnThrThrTyrTriPThrGlyThrThrTyrArgGlnIleThrAsnArgPhe 120
 QY 358 ACAGATAGGACGCCCGCTTCATTGAAGAGATCAGGATCTAATCGACGGCAAGGAAGA 417
 Db 121 ThrAspArgThrProValSerIleGluGluIleThrAspValIleAspAlaLysGlyArg 140
 QY 418 TGCATCATTAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGACAGGGAT 477
 Db 141 CysSerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaTyrAspGlyLysP 160
 QY 478 GCGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCTAGGGCA 537
 Db 161 AlaGlyGluLysGlnValLeuGlnProSerLysPheAsnThrProGluSerLysAla 180
 QY 538 TGGCACACGACTAATGAGACGTATACCGTGTGGGATCACCATGATATATCGAACGGGA 597
 Db 181 TrpHisThrThrAsnAspThrTyrThrValTyrGlySerProTyrValTyrArgThrGly 200
 QY 598 ACCFCCGTCATTTGTATAGTAGAAGAAATGATGCCGCTCTGTGTTCCGCTATTTCATAT 657
 Db 201 ThrSerValAsnCysIleValGluGluMetAspAlaArgAlaAlaPheProTyrSerTyr 220
 QY 658 TTTGCAATGGCCAAT 672
 Db 221 PheAlaMetAlaAsn 225

RESULT 11
 Q69406 PRELIMINARY; PRT; 865 AA.
 ID Q69406;
 AC Q69406;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GB homolog.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SB-1;
 RX MEDLINE=94233711; Pubmed=8178437;
 RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
 RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
 RT identification and expression by recombinant fowlpox viruses.";
 RL Virology 200:484-493(1994).
 DR EMBL; U01886; AAA19445.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 97236 MW; 04D4802FDC2E3948 CRC64;

```

Alignment Scores:
Pred. No.: 5,17e-96 Length: 865
Score: 990.50 Matches: 182
Percent Similarity: 89.78% Conservative: 20
Best Local Similarity: 80.89% Mismatches: 22
Query Match: 82.54% Indels: 1
DB: 12 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x Q69406 (1-865)
QY 1 ATGCACTATTATAGGCGGAATTCGATA---TTTTCCTTATAGTATTATCTATATGCTAGC 57
DB 1 MetAsnHisPheSerGlyLeuValProLeuPheSerValValPheValPhePhe 20
QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGCTCCAGTGG 117
DB 21 GlyArgValAlaArgAlaGlnAsnValThrSerArgGluAlaValSerValGlnLeu 40
QY 118 TCTGAGGAAGAGTCTACGTTTATCTTGTCCGCCACACAGTGGTTCACACCGTGTATCCGT 177
DB 41 SerGluAspGluSerSerPhePheLeuCysProAlaValGlyThrThrValValArg 60
QY 178 CTAGAACCGCGCGAAATGTCCCGAACCTAGAAAAGCCACGAGTGGGTGAAGAATC 237
DB 61 LeuGluProProArgLysCysProGluProLeuLysAlaThrGlnTrpGlyGluGlyIle 80
QY 238 GCGATATTATTAAGAGAAATATCAGTCCATATAAATTTAAAGTGACGCTTTATATAAA 297
DB 81 AlaileuPheLysGluAsnIleAsnProTyrLysPheLysValThrLeuTyrLys 100
QY 298 AATATCATTCAGACGACGATGACGGGACACATATAGACAGATCACTAATCATAT 357
DB 101 AsnValIleGlnThrThrTrpThrGlyThrTyrArgGlnIleThrAsnArgPhe 120
QY 358 ACAGATAGGACGCGCTTCCATTAAGAGATCAGGATCAGGATCACTAATCCGCAAGGA 417
DB 121 ThrAspArgThrProValSerIleGluGluIleThrAspValIleAspAlaLysGlyArg 140
QY 418 TGCTCATCTAAAGCAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGACGGAT 477
DB 141 CysSerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaTyrAspGlyAsp 160
QY 478 GCGGGAGAAAACAGTACTCTTAAACACCATCAAAATTCACACGCCGCGGATCAGGCA 537
DB 161 AlaGlyGluLysGlnValLeuLeuGlnProSerLysPheAsnThrProGluSerLysAla 180
QY 538 TGGCACACGACTAATGAGAGGTATACCGTGTGGGATCACCATGGATATATCAACGGGA 597
DB 181 TrpHisThrThrAsnAspThrThrValTrpGlySerProTrpValTyrArgThrGly 200
QY 598 ACCTCCGCTCAATTGTATAGTAGGAATGGATGCCGCTCTGTGTTTCGTTATTCATAT 657
DB 201 ThrSerValAsnCysIleValGluGluMetAspAlaArgAlaAlaLeuGlyTyrSerTyr 220
QY 658 TTTCGAATGCCAAT 672
DB 221 PheAlaMetAlaAsn 225

RESULT 12
Q90050 PRELIMINARY; PRT; 948 AA.
ID Q90050;
AC Q90050;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Herpes simplex virus glycoprotein B (gB) homolog.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID-10334;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-93090104; PubMed-1333759;
RA Maeda K., Horimoto T., Norimine J., Kawaguchi Y., Tomonaga K.,
RA Nikura M., Kai C., Takahashi E., Mikami T.;
RT "Identification and nucleotide sequence of a gene in feline
RT herpesvirus type 1 homologous to the herpes simplex virus gene
RT encoding the glycoprotein B.";
RL Arch. Virol. 127:387-397(1992).
DR EMBL; A49775; AAB24381.2; -;
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002048; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR PROSITE; PS00018; EF_HAND; 1.
SQ SEQUENCE 948 AA; 106233 MW; 71C28FB00B091325 CRC64;

Alignment Scores:
Pred. No.: 5,28e-62 Length: 948
Score: 670.00 Matches: 129
Percent Similarity: 59.43% Conservative: 36
Best Local Similarity: 45.91% Mismatches: 54
Query Match: 55.83% Indels: 60
DB: 12 Gaps: 3

US-09-147-052-3_COPY_1_672 (1-672) x Q90050 (1-948)
QY 7 TATTTAGCGGAATTCGATATTT----- 30
DB 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
DB 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCATATAT----- 51
DB 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 52 -----GGTACCAACTCATCTCCGAGTACCCAA 78
DB 82 ProArgArgThrValAlaThrProGluValGlyThrProProLysProThrThrAsp 101
QY 79 AATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTCTCTGAGGAAGAG----- 129
DB 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
QY 130 ---TCTACGTTTATCTTTGTCCCCACCAGTGGGTTCAACCGTGTACCGTCTAGAACCG 186
DB 122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
QY 187 CCGCGAAATGTCCCAACCTAGAAAAGCCACGAGTGGGTGAAGGAATCGGATATTA 246
DB 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
QY 247 TTTTAAAGAGAATATCAGTCCATATAAATTTAAAGTCACGCTTTATTTATAAAATATCAT 306
DB 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
QY 307 CACACGACGACATGGAGCGGACGACATATAGACAGATCACTAATCATATACAGATAG 366
DB 182 MetThrValTrpSerPheSerGlySerTyrTyrAlaValThrThrAsnArgTyrThrAspArg 201
QY 367 ACGCCGCTTTCATTAAGAGATCAGGATCTAATCGACGCGCAAGGAAGATGCTCATCT 426
DB 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSer 221
QY 427 AAAGCAGATACCTTGAACAATATATATGTTGAACCGTTTACAGGGATCGGGAGAA 486
DB 222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro 241
QY 487 AAACAAGTACTTCTAAACCATCAAAATTCACACGCGCGCAATCTAGGCGATGCGCACG 546
DB 242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGluSerArgGlyTrpHisThr 261

```

QY	412	GSAGATGTCATCTATAAGCAAGATACCTTTAGAAACAATGATATATGTTGAAGCGTTTGAC	471
Db	158	GlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyrGluPheThrAlaPheAsp	177
QY	472	AGGGATCGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACACGCCGAATCT	531
Db	178	LysAspGluAspProArgGluMetHisLeuLysProSerLysPheAsnThrProGlySer	197
QY	532	AGGCGATCGCACACGACTAATGACACGCTATACCGCTGTGGGATCACCATGGATATATACGA	591
Db	198	ArgGlyTrpHisThrThrAsnAspThyThrLysIleGlySerProGlyPheTyrArg	217
QY	592	ACGGGAACCTCCGTCATATTGATAGTAGAGGAATGGATGCCGCTGTGTTCCTCCGAT	651
Db	218	ThrGlyThrSerValAsnCysIleValIgluValAlaArgSerValTyrProTyr	237
QY	652	TCATATTTTGCATGGCC	669
Db	238	AspSerPheGlyIleSer	243
RESULT 14			
Q91SD0	Q91SD0	PRELIMINARY;	PRT; 879 AA.
AC	Q91SD0;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Glycoprotein B.		
GS	Canine herpesvirus.		
OC	Viruses; dsDNA viruses, no RNA stage: Herpesviridae;		
OC	Alphaherpesvirinae.		
OX	NCBI_TaxID=37110;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SFRAN-Australian;		
RA	Reubel G.H., Pekin J., Webb-Wagg K., Hardy C.M.;		
RT	"Nucleotide sequence of glycoprotein genes B, D, G, H and I, thymidine		
RT	kinase and protein kinase genes and gene homolog UL 24 of an		
RL	Australian isolate of canine herpesvirus."		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF361073; AAK51052.1; -;		
DR	InterPro: IPR000234; Glycoprot_B.		
DR	Pfam: PF00606; Glycoprotein_B; 1.		
DR	ProDom: PD000693; Glycoprot_B; 1.		
SO	SEQUENCE 879 AA; 99584 MW; D37E2C586D7E59A1		
	CRC64;		

```

Alignment Scores:
pred. No.:      6,73e-61      Length:      879
Score:          659.50        Matches:     125
Percent Similarity: 68.10%    Conservative: 33
Best Local Similarity: 53.88% Mismatches:    57
Query Match:      54.96%     Indels:       17
DB:              12         Gaps:        2

US-09-147-052-3_COPY_1_672 (1-672) x Q91SD0 (1-879)

QY      25 ATATTTTCCTATAGTATTCTA----- 48
Db      8 ILePhePheIleIeIyThrLeuIleIeCysAspProThrThrProGluSerThrIle 27
QY      49 -----TATGTCAGAACTCATCTCCGAGTACCCAAAATGTGCATCAGAGAA 96
Db      28 AsnProLeuAsnHISHisAsnLeuSerThrProLysProThrSerAspSpleargGlu 47
QY      97 GTTGTTTCGAGCGTCCAGTTG---TCTGGAAGAAGTCTACGTTTTTATCTTTGTCCCCCA 153
Db      48 ILeuAArgGluSerGlnIleGluSerAspSphThrSerThrPheTyrMetCysProPro 67
QY      154 CCAGTGGGTTCACCGTGTACCGTCTAGAACCGCGCGAAATGTCGCGAACCCTAGAAAA 213
Db      68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrIysLeu 87

```

QY	214	GCACCCGAGTGGGTGAAGGAATCCGATATATTATTTAAAGAGAATATCATGTCATATAAA	273
Db	88	GlyLysAsnPheThrGluGlyLeuAlaValIlePheLysGluAsnIleSerProTyrLys	107
QY	274	TTTAAAGTGACGCTTTATTATATAAAATATCATTCAGACGACGACATGGAGGGGACGACA	333
Db	108	PheLysAlaAsnIleTyrTyrLysAsnIleIleThrValTrpSerGlySerThr	127
QY	334	TATGACAGATCACTAAATCGATATACAGATAGGACGCCCTTTCCATTGAAGAGATCAGC	393
Db	128	TyrAlaValIleThrAsnArgTy-ThrAspArgValProIleGlyValProGluIleThr	147
QY	394	GATCTAATTCGACGGCAAGAGAGATGCTCATCTATAACCAACAGATACCTTAGAACAAATGTA	453
Db	148	GluLeuIleAspArgArgGlyMeCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr	167
QY	454	TATGTTGAAGCGTTTGACAGGAGTCCGGGAGAAACAACTACTTCTAAACCATCAAAA	513
Db	168	GluPheThrAlaPheAspLysAspLysAspProArgGluValHisLeuLysProSerLys	187
QY	514	TTACACAGCCCGGAATCTAGGCGATGGCACACGACTAATGAGAGGTATACCGTGTGGGA	573
Db	188	PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly	207
QY	574	TCACCATGGATATATCGAACGGGAACCTCCGCTCAATTGTATAGTAGAGAAATGGATGCC	633
Db	208	GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla	227
QY	634	CGCTCTGTGTTCCGATATTCATATTTGCAATGCC	669
Db	228	ArgSerValTyrProTyrAspSerPheAlaIleSer	239

RESULT 15

Q98VM4

PRELIMINARY; PRT; 868 AA.

AC Q98VM4;

01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

ORF31.

DE Human herpesvirus 3.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

ON NCBI_TaxID=10335;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ZV1-LAX1, and ZV2-Oka;

RX MEDLINE=21109116; PubMed=11162813;

RA Faga B., Maury W., Bruckner D.A., Grose C.;

RT "Identification and Mapping of Single Nucleotide Polymorphisms in the

RL Varicella-Zoster Virus Genome.,"

RT Virology 280:1-6(2001).

DR EMBL; AY016451; AAK19964.1; -.

DR EMBL; AY016445; AAK19938.1; -.

DR InterPro; IPR000234; Glycoprot.B.

DR InterPro; IPR000531; TonB_box.

DR Pfam; PF00606; Glycoprotein.B; 1.

DR ProDom; PD000693; Glycoprot.B; 1.

DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.

SQ SEQUENCE 868 AA; 98062 MW; 536B91E101A16AD1 CRC64;

Alignment Scores:	
Pred. No.:	4.5e-56
Score:	614.00
Percent Similarity:	75.96%
Best Local Similarity:	54.61%
Query Match:	51.17%
DB:	12
Length:	868
Matches:	108
Conservative:	36
Mismatches:	49
Indels:	0
Gaps:	0

US-09-147-052-3_COPY_1_672 (1-672) x Q98VM4 (1-868)

QY 91 AGAGAAGTTCTTTCGACGGTCCAGTGTCTGAGGAAGACTCTACGTTTATCTTTGTCCC 150

Db	41	ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro	60
Qy	151	CCACGAGGGTTCAACCGTCATCGCTAGAACCGCGGAAATGTGCCGAACCTAGA	210
Db	61	ProProThrGlySerThrIleValArgLeuGluProProArgThrCysProAspTyrHis	80
Qy	211	AAAGCCACCGAGTGGGGTGAAGGAATCCGATATATTATTAAGAGAATATCATCAGTCCATAT	270
Db	81	LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaIaTyr	100
Qy	271	AAATTTAAAGTCAGCCTTATATATAAATATCATTCAGACGACGACATGGACGGGGACG	330
Db	101	LysPheLysAlaThrValTyrLysAspValIleValSerThrAlaThrPalaGlySer	120
Qy	331	ACATATAGACAGATCACTAAATCGATATACAGATAGGACGCCGTTTCATTCATTAAGACGATC	390
Db	121	SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle	140
Qy	391	ACGGATCTAATCGACGGCGAAAGAGATGCTCATCTAAAGCAAGATACATCTTAGAACCAAT	450
Db	141	ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn	160
Qy	451	GTATATGTTGAAGCGTTTGACAGGGATCGGGAGGAAACAAACAGTACTTCTAAACCAATCA	510
Db	161	HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer	180
Qy	511	AAATTCACACGCCCGAATCTAGGCGATGGACACGACTATGACACGCTATACCGGTGG	570
Db	181	LysTyrAsnSerValGlySerLysAlaThrPheIleThrAsnAspThrTyrMetValAla	200
Qy	571	GGATCACCATCGATATCGAACGGGAACCTCCGTCAATTGTATAGTAGAGAAATGGAT	630
Db	201	GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluValGlu	220
Qy	631	GCCCGCTCTGTTCCTCGATTCATATTTTCGATGGCC	669
Db	221	AlaArgSerIlePheProTyrAspSerPheGlyLeuSer	233

Search completed: October 8, 2003, 18:12:31
Job time : 66.5355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:33:17 ; Search time 10.8306 Seconds
(without alignments)
5835.661 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200

Sequence: 1 atgcaactatttagcgga.....catattttgcattgccaat 672

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlpl
-Q/cgn2_1/USPTO_spool/US09147052/runat_08102003_154340_29835/app_query.fasta_1.5980
-DB=SwissProt_41 -OPM=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_140@runat_08102003_154340_29835 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1200	100.0	865	1 VGLB_HSVMD	P18538 marek's dis
2	606	50.5	868	1 VGLB_PVVD	P09257 varicella-z
3	605	50.4	913	1 VGLB_PVIV	P08355 pseudorabie
4	583	48.6	928	1 VGLB_HSVBP	P17471 bovine herp
5	564.5	47.0	932	1 VGLB_HSVBC	P12640 bovine herp
6	563.5	47.0	979	1 VGLB_HSVEL	P25218 equine herp
7	563.5	47.0	980	1 VGLB_HSVBA	P18551 equine herp
8	563.5	47.0	980	1 VGLB_HSVBE	P28922 equine herp
9	559.5	46.6	919	1 VGLB_HSVFA	P17472 equine herp
10	529.5	44.1	933	1 VGLB_HSVAI	Q04463 herpesvirs
11	526	43.8	920	1 VGLB_HSVSM	Q04464 herpesvirs
12	525.5	43.8	980	1 VGLB_HSVBI	P18050 equine herp
13	504	42.0	885	1 VGLB_HSVS2	P24994 herpes smp
14	504	42.0	903	1 VGLB_HSVIF	P06436 herpes smp
15	504	42.0	904	1 VGLB_HSV11	P10211 herpes smp
16	504	42.0	904	1 VGLB_HSV1K	P06437 herpes smp
17	504	42.0	904	1 VGLB_HSV1P	P06565 herpes smp
18	500	41.7	904	1 VGLB_HSV23	P06763 herpes smp

19	499	41.6	917	1 VGLB_HSVB2	P12641 bovine herp
20	498	41.5	904	1 VGLB_HSV2H	P08666 herpes smp
21	426.5	35.5	883	1 VGLB_ILTVS	P27415 infectious
22	422	35.2	873	1 VGLB_ILTV6	Q02409 infectious
23	422	35.2	883	1 VGLB_ILTVT	P24904 infectious
24	219	18.2	944	1 VGLB_HSVT2	Q9wrl5 herpesvirs
25	211.5	17.6	808	1 VGLB_HSVSA	P24905 herpesvirs
26	211.5	17.6	830	1 VGLB_HSV62	P36320 human herpe
27	209.5	17.5	822	1 VGLB_HSV7J	P52352 human herpe
28	209.5	17.5	830	1 VGLB_HSV6G	P36319 human herpe
29	209.5	17.5	830	1 VGLB_HSV6U	P28864 human herpe
30	202.5	16.9	901	1 VGLB_GPCMV	Q69024 guinea pig
31	202.5	16.9	907	1 VGLB_HCMVT	P13201 human cytom
32	201.5	16.8	857	1 VGLB_EBV	P03188 Epstein-Bar
33	194.5	16.2	906	1 VGLB_HCMVA	P06473 human cytom
34	189.5	15.8	854	1 VGLB_RHCM6	P89053 rhesus cyto
35	180	15.0	928	1 VGLB_MCMVS	P27171 murine cyto
36	89	7.4	848	1 VGLB_HSVB	P09101 equine herp
37	88.5	7.4	1185	1 MAPX_DROME	P23226 drosophila
38	88	7.3	416	1 LMP1_HUMAN	P11279 homo sapien
39	87.5	7.3	306	1 DIMH_DROME	Q9vaq5 drosophila
40	86.5	7.2	262	1 BCSX_ACEXY	Q9wa69 acetobacter
41	86	7.2	536	1 CCA4_DROME	Q9ve00 drosophila
42	85	7.1	3415	1 P0LG_POWVL	Q04538 t genome po
43	84	7.0	1531	1 YQ38_CAEEL	Q09459 caenorhabdi
44	84	7.0	3726	1 TRX_DROME	P20859 drosophila
45	82.5	6.9	467	1 SIL5_MOUSE	Q91y57 mus musculu

ALIGNMENTS

RESULT 1

ID	VGLB_HSVMD	STANDARD;	PRT;	865 AA.
AC	P18538;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	GB.			
OS	Marek's disease herpesvirus (strain RB-1B) (MDHV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_Taxid=33707;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89293086; PubMed=2544666;			
RA	Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;			
RT	"Nucleotide sequence and characterization of the Marek's disease			
RT	virus homologue of glycoprotein B of herpes simplex virus.";			
RL	J. Gen. Virol. 70:1789-1804(1989).			
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; D13713; BAA02866.1; -			
DR	InterPro; IPR000234; Glycoprot_B.			
DR	Pfam; PF00606; Glycoprotein_B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			
KW	Signal; Glycoprotein; Transmembrane.			
FT	SIGNAL 1 21 POTENTIAL.			
FT	CHAIN 22 865 GLYCOPROTEIN B.			
FT	DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 683 700 POTENTIAL.			
FT	TRANSMEM 709 729 POTENTIAL.			
FT	TRANSMEM 732 752 POTENTIAL.			
FT	DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).			

FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 98091 MW; B30E93C1AC65C6C3 CRC64;

Alignment Scores:
 Pred. No.: 6,17e-113 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVMD (1-865)

QY 1 ATGCACTATTAGCGGAATTCATATTTTCTTACTATCTTCTATATCTATGTCAGCAAC 60
 Db 1 MethisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCACCGTGATCCGCTA 180
 Db 41 GluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAACCGCGCGGAAATGTCGCAACCTAGAAAGCCAGGAGTGGGTGAGGAATCGCG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
 QY 241 ATATTATTAAAGAGAATACAGTCCATATAAATTTAAAGTGAGCGCTTTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
 QY 301 ATCAATTCAGAGCAGATGCGGCGGACGACATATAGACAGATCACTAATCGATATACA 360
 Db 101 IleIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120
 QY 361 GATAGGACGCGCTTTCCATTAAGAGATCACGGATCTAATCGACGCGCAAGGAATGTC 420
 Db 121 AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
 QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGATATGTTGAAGCGTTTGACAGGATCGC 480
 Db 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAspAla 160
 QY 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATCAACGCGCCGCAATCTAGGCGATGG 540
 Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
 QY 541 CACAGCACTATGAGAGTATACCGTGTGGGATCACCATGATATATCAACGCGGAACC 600
 Db 181 HisThrThrAsnGluThrTyrThrValTrpLysSerProTrpIleTyrArgThrGlyThr 200
 QY 601 TCCGTCATATGATAGAGAAATGATGCGCGCTGCTGTTCCGTTATCATATTTT 660
 Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
 QY 661 GCAATGGCCAAT 672
 Db 221 AlaMetAlaAsn 224

RESULT 2

VGLB_VZVD

ID VGLB_VZVD

AC P09257; STANDARD; PRT; 868 AA.

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycoprotein B precursor (Glycoprotein II).
 GN 31.
 OS Varicella-zoster virus (strain Dumas) (VZV).
 ON Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86306657; PubMed=3018124;
 RA Davison A.J., Scott J.E.;
 RL "The complete DNA sequence of varicella-zoster virus.";
 J. Gen. Virol. 67:1759-1816(1986).
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; X04370; CAA27914.1; -.
 PIR; E27214; VGBE31.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 KW Glycoprotein; transmembrane; signal.
 FT SIGNAL 1 ?
 FT CHAIN ? 868 GLYCOPROTEIN B.
 FT DOMAIN ? 679 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 680 695 POTENTIAL.
 FT TRANSMEM 701 720 POTENTIAL.
 FT TRANSMEM 724 744 POTENTIAL.
 FT DOMAIN 745 868 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 868 AA; 98066 MW; 58B8E1346F4FF902 CRC64;

Alignment Scores:

Pred. No.: 5,3e-53 Length: 868
 Score: 606.00 Matches: 107
 Percent Similarity: 74.09% Conservatives: 36
 Best Local Similarity: 55.44% Mismatches: 50
 Query Match: 50.50% Indels: 0
 DB: 1 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_VZVD (1-868)

QY 91 AGAAGATTGTCGAGCGTCCAGTTCGTGAGGAAGAGTCTAGCTTTTATCTTTGTC 150
 Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
 QY 151 CCACCACTGGTCAACCGTCATCCGCTAGACCGCGCGGAAATCTCCGAACCTAGA 210
 Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
 QY 211 AAGCCACCGAGTGGGTGAAGGAATCGCATATATTATTAAAGAGAATATCAGTCCATAT 270
 Db 81 LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaTyr 100
 QY 271 AAATTTAAAGTGACGCTTTATTTAAATATCATTCAGAGGACGACATGACGGGACG 330
 Db 101 LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTyrPheAlaGlySer 120
 QY 331 ACATATAGACAGATCACTAATCGATATACAGATAGGACCGCGCTTCCATTCAGAGATC 390
 Db 121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140

QY 391 ACGGATCTAAGCAGCGCAAGAGATGCTCATCTAAGCAAGATACCTTAGAACAT 450
 Db 141 ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrThrTyrValArgAsnAsn 160
 QY 451 GTATATGTTGAAGCGTTTACAGGGATGCGGAGAGAAACAAAGTACTTCTAAACCATCA 510
 Db 161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
 QY 511 AAATTCAACACGCCGAATCTAGGGATGGCACAGCACTAAGTAAAGTATACCGTGG 570
 Db 181 LysTyrAsnSerValGlySerLysAlaThrPheHisThrThrAsnAspThrTyrMetValAla 200
 QY 571 GGATCACCATGATATATCATCAAGCGGACCTCCGTCGAATGTATAGTACAGGAATGGAT 630
 Db 201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluValGlu 220
 QY 631 GCCCGCTCTGTTCCGTCATTCATATTTGCAATGGCC 669
 Db 221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSer 233
 RESULT 3
 VGLB_PRVIF STANDARD; PRT; 913 AA.
 AC P0835;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein gII precursor.
 OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_Taxid:31523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87284141; PubMed=3039163;
 RA Robbins A.K., Dorsey D.J., Wathen M.W., Whealy M.E., Gold C.,
 RA Watson R.J., Holland L.E., Weed S.D., Levine M., Glorioso J.C.,
 RA Enquist L.W.;
 RT "Pseudorabies virus gII gene is closely related to the gB
 glycoprotein gene of herpes simplex virus.";
 RL J. Virol. 61:2691-2701(1987).
 RN [2]
 RP SEQUENCE OF 847-913 FROM N.A.
 RX MEDLINE=89279298; PubMed=2543777;
 RA Simon A., Mettenleiter T.C., Rziha H.J.;
 RT "Pseudorabies virus displays variable numbers of a repeat unit
 adjacent to the 3' end of the glycoprotein gII gene.";
 RL J. Gen. Virol. 70:1239-1246(1989).
 RN [3]
 RP EXPORT PATHWAY.
 RX MEDLINE=90219190; PubMed=2157862;
 RA Whealy M.E., Robbins A.K., Enquist L.W.;
 RT "The export pathway of the pseudorabies virus gB homolog gII involves
 oligomer formation in the endoplasmic reticulum and protease
 processing in the Golgi apparatus.";
 RL J. Virol. 64:1946-1955(1990).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M17321; AAA47465.1; -;
 DR EMBL; D00464; BAA00359.1; -;
 DR PIR; A29159; VGBEPS.
 DR InterPro; IPR000234; Glycoprot.B.
 DR Pfam; PF00606; Glycoprotein.B; 1.
 DR ProDom; PD000693; Glycoprot.B; 1.

KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 40
 FT CHAIN 41 913 GLYCOPROTEIN GII.
 FT DOMAIN 41 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 819 POTENTIAL.
 FT DOMAIN 820 913 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 516 516 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 913 AA; 100234 MW; 5D560D235E856437 CRC64;
 Alignment Scores:
 Pred. No.: 6.77e-53 Length: 913
 Score: 605.00 Matches: 112
 Percent Similarity: 72.12% Conservative: 38
 Best Local Similarity: 53.85% Mismatches: 54
 Query Match: 50.42% Indels: 4
 DB: 1 Gaps: 2
 US-09-147-052-3_COPY_1_672 (1-672) x VGLB_PRVIF (1-913)
 QY 52 GGTCAGACTCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAAGTGTTCG 105
 Db 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
 QY 106 ACGGTCCAGTGTCTGAGGAAGAGTCTAGCTTTTATCTTGTGCCCCCACCAGTGGTTCA 165
 Db 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
 QY 166 ACCGTGATCCGCTAGAACCGCGCCGAAATGTCGCGAAACCTAGAGAAAGCCACGATGG 225
 Db 133 ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
 QY 226 GGTGAAGGAATCGCATATATTAAAGAGATATATCAGTCCATATATAAATTTAAAGTGACG 285
 Db 153 ThrGluGlyIleAlaValLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
 QY 286 CTTTATTATAAAATATCATTCAGACGACATGCGGAGGACGACATATAGACAGATC 345
 Db 173 IleTyrTyrLysAsnValIleValThrValTrpSerGlySerThrTyrAlaAlaIle 192
 QY 346 ACTAATCGATATACAGATAGAGCGCCGTCATTTCATTGAAGAGATCAGCATCTAATCGAC 405
 Db 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
 QY 406 GGCAAAGGAAGATGCTCATCTTAAGCAAGATACCTTAGAACAAATGTATATGTTGAAGCG 465
 Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
 QY 466 TTTCAGAGGATGCGGGAGAGAAACAAAGTACTTCTAAACCATCAAAATTCACACGCCCC 525
 Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
 QY 526 GAATCAGGCGCATGGCACACGACTAATGAGACGTATACCGTGTGGGAGTACCATCGATA 585
 Db 253 GlyThrArgGlyTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
 QY 586 TATCGAACGGGAACCTCCGTCATTTGTATGTATAGTAGAGAAATGGATCCCGCTGTGTGT 645
 Db 273 TyrHisThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292
 QY 646 CCGTATTTCATATTTTCATGCGCC 669
 Db 293 ProTyrAspSerPheAlaLeuSer 300
 RESULT 4
 VGLB_HSVBP
 ID VGLB_HSVBP STANDARD; PRT; 928 AA.
 AC P17471;

	01-AUG-1990	(Rel. 15,	Created)
Dt	01-AUG-1990	(Rel. 15,	Last sequence update)
Df	16-OCT-2001	(Rel. 40,	Last annotation update)
De	Glycoprotein I precursor	(Glycoprotein GVP-6)	(Glycoprotein LIA)
DE	(Glycoprotein 16)	(Glycoprotein GL30).	
GN	GI.		
OS	Bovine herpesvirus type 1	(strain PB-2).	
OC	Viruses; dsDNA viruses, no RNA stage;	Herpesviridae;	
Ox	Alphaherpesvirinae;	Variellovirus.	
OX	NCBI_TaxID=10324;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=89020821; Pubmed=2845660;		
KX	Misra V., Nelson R., Smith M.;		
RT	*Sequence of a bovine herpesvirus type-1 glycoprotein gene that is homologous to the herpes simplex virus gene for the glycoprotein gB.*;		
RL	Hirology 166:542-549(1988).		
RL	-1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.		
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensee@lsb-sib.ch).		
CC	EMBL; M23257; AAA46013.1; ..		
DR	PTR; A31166; VGEBEG.		
DR	InterPro: IPR000234; Glycoprot_B.		
DR	Fram; PF00606; Glycoproteins_B; 1.		
DR	ProdDom; PD000693; Glycoprot_B; 1.		
KW	Glycoprotein; Transmembrane; Signal.		
FT	SIGNAL 1..67		
FT	CHAIN 68..928		
FT	DOMAIN 68..780		
FT	TRANSMEM 781..801		
FT	TRANSMEM 804..824		
FT	DOMAIN 825..928		
FT	CARBOHYD 105..105		
FT	CARBOHYD 153..153		
FT	CARBOHYD 442..442		
FT	CARBOHYD 484..484		
FT	CARBOHYD 579..579		
FT	CARBOHYD 637..637		
FT	CARBOHYD 703..703		
FT	CONFLICT 409..409		
FT	CONFLICT 673..673		
FSQ	SEQUENCE 928 AA; 10217 MW; B47982224FCDF69D CRC64;		

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
52	1,13e-50	583.00	107
Percent Similarity:		68.45%	
Best Local Similarity:		51.94%	
Query Match:		48.58%	
DB:		1	
			Conservative: 35 Mismatched: 64 Indels: 0 Gaps: 0

```
US-09-147-052-3_COPY_1_672 (1-672) x VLBU_HSVBP (1-928)
Pred. No.:    Length:   Score:   Matches:
52            1,13e-50  583.00     107
Percent Similarity: 68.45%
Best Local Similarity: 51.94%
Query Match:      48.58%
DB:               1
                Conservative: 35
                  Mismatched: 64
                    Indels: 0
                      Gaps: 0

QY          52 GTACGAAGTCATCTCCCGTAGTACCCAAAATGCACATCAAGAAGACTGTGGTTTGAGCGTC 111
Db           ||| : ::::: |
              97 GlyAspApsAlaSerProaspasnSerThspValargalaLeuArgLeuAla 116

QY         112 CAGTTGTCGAGAGAGCHTACGTTTTATTCTTTGCCCCCCCAGTGGGTTCAACCGTG 171
Db          ||| : ::||| 
              117 GlnAlaLaagLygluaSnSerArgPhPeVaCyProProSerGlyAlathrVal 136

QY        172 ATCCGCTAGAACCGCCGAAAANTGTCGCCGACCTAGAAAAAGCCCAGTGCGGGTGA 231
Db         :::|::::| ||| ||| |::::|
              137 Vala:gLeualaPaalaaarProcysProGlutyrGlyLeuGlyArgasnyThrGlu 156
```

FT CHAIN 58 932 GLYCOPROTEIN I.
 FT DOMAIN 68 758 EXTRACELLULAR.
 FT TRANSMEM 759 827 POTENTIAL.
 FT DOMAIN 828 932 CYTOPLASMIC.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;

Alignment Scores:
 Pred. No.: 8,3e-49 Length: 932
 Score: 564.50 Matches: 106
 Percent Similarity: 67.96% Conservative: 34
 Best Local Similarity: 51.46% Mismatches: 65
 Query Match: 47.04% Indels: 1
 DB: 1 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVBC (1-932)

QY 52 GTAGAACTCATCTCCGAGTAGTACCAAAATGTGACATCAAGAGAGAGTGTGTTTCGAGCGTC 111
 DB 97 GlyAspAspAlaIaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116

QY 112 CAGTTGCTCTGAGGAGAGTCTACGTTTATCTTTCTCCACACAGTGGTTCACACCGTG 171
 DB 117 GlnAlaAlaGlyGluAsnSerArgPheValCysProProSerGlyAlaThrVal 136

QY 172 ATCCGCTAGAACCCCGCGGAAATGTCGCCAGCTAGAAAGCCACCGAGTGGGTGAA 231
 DB 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156

QY 232 GGAATCGCGATATTTAAAGAGATATCATGCTCATATAAATTTAAAGTACGCTTAT 291
 DB 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIle--- 175

QY 292 TATAAAATATCATTCAGACGACGATCGGAGCGGACGACATATAGACATCACTAAT 351
 DB 176 TyrLysAsnValIleValThrThrThrTrpAlaGlySerThrTyrAlaAlaIleThrAsn 195

QY 352 CATATACAGATAGACGCCCGCTTCCATTGAAGATCAGGATCAATCAATCGACGCA 411
 DB 196 GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLysLys 215

QY 412 GGAAGATGCTCATCTAAAGCAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTAC 471
 DB 216 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValAlaAlaPheAsp 235

QY 472 AGGATGCGGGAGAAAACAAGTACTTCTAAACCAATCAAAATCAACACGCCGAATCT 531
 DB 236 ArgAspAspProTrpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 255

QY 532 AGGGATGCCACACCACTAATCAGACGATACCGTGTGGGATCACCATTGATATATCGA 591
 DB 256 ArgGlyTrpHisThrAspAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 275

QY 592 ACGGGAACCTCGTCAATGTATAGTAGGAAATGGATGCCGCTCTGTGTTCCGAT 651
 DB 276 ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 295

QY 652 TCATATTTTGCATGGCC 669
 DB 296 AspSerPheAlaLeuSer 301

RESULT 6
 VGLB_HSVEL
 ID VGLB_HSVEL STANDARD; PRT; 979 AA.
 AC P25218;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33.
 OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90236317; PubMed=1692002;
 RA Guo P.;
 RT "Characterization of the gene and an antigenic determinant of equine
 herpesvirus type-1 glycoprotein 14 with homology to gB-equivalent
 glycoproteins of other herpesviruses.";
 RL Gene 87:249-255(1990).
 CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34861; AAA46086.1; -;
 DR PIR; JH0109; JH0109.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86 POTENTIAL.
 FT CHAIN 87 979 GLYCOPROTEIN B.
 FT DOMAIN 87 851 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 852 869 POTENTIAL.
 FT CARBOHYD 870 979 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 979 AA; 110331 MW; 9A19866B791C5B36 CRC64;

Alignment Scores:
 Pred. No.: 1,06e-48 Length: 979
 Score: 563.50 Matches: 113
 Percent Similarity: 59.36% Conservative: 36
 Best Local Similarity: 45.02% Mismatches: 69
 Query Match: 46.96% Indels: 33
 DB: 1 Gaps: 4

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVEL (1-979)

QY 13 AGCGGGAATTCATATTTTTCCTTATAGTATATCTATATGTT----- 54
 DB 66 ArgValGlyThrIleValLeuAlaCysLeuLeuLeuPheGlySerCysValValArgAla 85

QY 55 -----ACGAATCATCTCCGAGTACCCAAATGTGACATCA----- 90
 DB 86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105

QY 91 -----AGAGAGTGTGTT 102
 DB 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125

QY 103 TCGAGGCTCCAGTGTCTGAGGAGAGTCTACGTTTTTATCTTTTCTCCACCGAGGCT 162
 DB 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145

QY 163 TCAACCGTGATCCGTAGAACCGCGGAAATAGTCCGAAACCTAGAAACCCACCGAG 222
 DB 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
 QY 223 TGGGGTGAAGGAATCGCGATATATTAAAGAGATATCATGCTCCATATAAATTTAAAGTG 282
 DB 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla 185
 QY 283 ACGTTTATATAAATATCATTCAGACGACGACATGAGCGGGGAGCAGCATATAGACAG 342
 DB 186 AsnValTyrTyrLysAspIleValValThrArgValTyrLysGlyTyrSerHisThrSer 205
 QY 343 ATCACTAATCATATACAGTACGACCGCGCTTCCATTTGAGAGATCAGCGATCTAATC 402
 DB 206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuLeu 225
 QY 403 GACGGCAAGGAAGATGCTCATCTAAAGCAGATACCTTGAACAACATGTATATGTTGAA 462
 DB 226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMetHisHis 245
 QY 463 GCGTTTCACAGGATGCGGGAGAAAACACAGTACTTCTTAAACCATCAAAATTCACACG 522
 DB 246 AlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
 QY 523 CCCGAATCTAGGCGATGCGACGACTAATGAGACG-----TATACCGTGTGGGATCA 576
 DB 266 ProGlyAlaArgAlaTyrGlnThrThrAsnAspThrThrSerTyrValGlyTyr--Met 284
 QY 577 CCATGGATATATCGAACGGGAACCTCGTCAATGTATAGTAGAGAAATGCGCCGC 636
 DB 285 ProTyrArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAlaArg 304
 QY 637 TCTGTGTTCCGTATTCATATTTTGCATGGCC 669
 DB 305 SerValTyrProTyrAspSerPheAlaLeuSer 315
 RESULT 7
 VGLB_HSVEA STANDARD; PRF: 980 AA.
 ID P18531;
 AC 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33
 OS Equine herpesvirus type 1 (strain Abi) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10328;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bonass W.A., Elton D.M., Stocks J.M., Killington R.A.,
 RA Meredith D.M., Halliburton I.W.;
 RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M36298; AAA46068.1; --
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 DR KEGG; Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 980 GLYCOPROTEIN B.
 FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 853 870 POTENTIAL.

FT DOMAIN 871 980
 FT CARBOHYD 165
 FT CARBOHYD 275
 FT CARBOHYD 380
 FT CARBOHYD 423
 FT CARBOHYD 497
 FT CARBOHYD 514
 FT CARBOHYD 515
 FT CARBOHYD 560
 FT CARBOHYD 727
 FT CARBOHYD 749
 SQ SEQUENCE 980 AA; 109736 MW; 264D273CED70E5A1 CRC64;
 Alignment Scores:
 Pred. No.: 1.06e-48 Length: 980
 Score: 563.50 Matches: 113
 Percent Similarity: 59.36% Conservative: 36
 Best Local Similarity: 45.02% Mismatches: 69
 Query Match: 46.96% Indels: 33
 Db: 1 Gaps: 4
 US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVEA (1-980)
 QY 13 AGCGGAATTCATATTTTCCCTTATAGTATTCTATATGCT----- 54
 DB 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
 QY 55 -----ACGAACTCATCTCCGAGTACCACCAAAATGTGACATCA----- 90
 DB 86 ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis 105
 QY 91 -----AGAGAAGTGTGTT 102
 DB 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
 QY 103 TCGAGCGTCCAGTTGCTGTGAGGAAGAGTCTACGTTTCTTCTCTCCGCCACCGAGTGGGT 162
 DB 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145
 QY 163 TCAACCGTGATCCGTAGAACCGCGGAAATAGTCCCGAACCTAGAAACCCACCGAG 222
 DB 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
 QY 223 TGGGGTGAAGGAATCGCGATATATTAAAGAGATATCATGCTCCATATAAATTTAAAGTG 282
 DB 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla 185
 QY 283 ACGTTTATATAAATATCATTCAGACGACGACATGAGCGGGGAGCAGCATATAGACAG 342
 DB 186 AsnValTyrTyrLysAspIleValValThrArgValTyrLysGlyTyrSerHisThrSer 205
 QY 343 ATCACTAATCATATACAGTACGACCGCGCTTCCATTTGAGAGATCAGCGATCTAATC 402
 DB 206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuLeu 225
 QY 403 GACGGCAAGGAAGATGCTCATCTAAAGCAGATACCTTGAACAACATGTATATGTTGAA 462
 DB 226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMetHisHis 245
 QY 463 GCGTTTCACAGGATGCGGGAGAAAACACAGTACTTCTTAAACCATCAAAATTCACACG 522
 DB 246 AlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
 QY 523 CCCGAATCTAGGCGATGCGACGACTAATGAGACG-----TATACCGTGTGGGATCA 576
 DB 266 ProGlyAlaArgAlaTyrGlnThrThrAsnAspThrThrSerTyrValGlyTyr--Met 284
 QY 577 CCATGGATATATCGAACGGGAACCTCGTCAATGTATAGTAGAGAAATGCGCCGC 636
 DB 285 ProTyrArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAlaArg 304
 QY 637 TCTGTGTTCCGTATTCATATTTTGCATGGCC 669
 DB 305 SerValTyrProTyrAspSerPheAlaLeuSer 315

Db 305 SerValTyProTyAspSerPheAlaLeuSer 315

RESULT 8

VGLB_HSVB
ID VGLB_HSVB STANDARD; PRT; 980 AA.
AC P28922;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520, 10330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AB4P;
RX MEDLINE=922595566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1";
RL Virology 189:304-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Kentucky D;
RX MEDLINE=90219246; PubMed=2157895;
RA Guo P.X., Goebel S., Perkus M.E., Taylor J., Norton E., Allen G.,
RA Lanquet B., Desmettre P., Paoletti E.;
RT "Coexpression by vaccinia virus recombinants of equine herpesvirus 1
glycoproteins gp13 and gp14 results in potentiated immunity";
RL J. Virol. 64:2399-2406(1990).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M86664; AAB02468.1; -
DR EMBL; M35145; AAR46067.1; -
DR PIR; G36798; VGBEC6.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 980
FT DOMAIN 87 852
FT TRANSMEM 853 870
FT TRANSMEM 853 870
FT DOMAIN 871 980
FT CARBOHYD 165 165
FT CARBOHYD 275 275
FT CARBOHYD 380 380
FT CARBOHYD 423 423
FT CARBOHYD 497 497
FT CARBOHYD 514 514
FT CARBOHYD 515 515
FT CARBOHYD 560 560
FT CARBOHYD 727 727
FT CARBOHYD 749 749
SQ SEQUENCE 980 AA; 109805 MW; A6DDFA8CA5550FF5 CRC64;

Alignment Scores:

Pred. No.: 1,06e-48 Length: 980
Score: 563.50 Matches: 113
Percent Similarity: 59.36% Conservative: 36
Best Local Similarity: 45.02% Mismatches: 69
Query Match: 46.96% Indels: 33
DB: 1 Gaps: 4

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVB (1-980)
QY 13 AGCGGAATTCATATTTTCTCTTATAGTTATTTCTATATGTT----- 54
DB 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
QY 55 -----ACGAATCATCTCCGAGTACCCAAAATGTGACATCA----- 90
DB 86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105
QY 91 -----AGAGAACTTGT-----AGAGAACTTGT 102
DB 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY 103 TCGAGCGTCCAGTCTGTGAGGAAGAGCTACGTTTATCTTGTCTCCGCCACCTGGGT 162
DB 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheThrThrCysProProThrGly 145
QY 163 TCACCGTGATCGGTCTAGAACCGCGGGAATATCTCCGAACCTAGAAAACCCACCGAG 222
DB 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGTGAGGATCGCATATTTTAAAGAGATATCATGTCCTCATTAATTTAAAGTG 282
DB 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyLysPheArgAla 185
QY 283 AGCGTTTATATAAAATATCATTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 342
DB 186 AsnValTyTyLysAspIleValValThrArgValThrLysGlyTySerHisThrSer 205
QY 343 ATCACTAATCGATATACAGTAGAGACGCGCGTTCCATGAGAGATCAGGATCTAATC 402
DB 206 LeuSerAspArgTyAsnAspArgValProValSerValGluGluIlePheGlyLeu 225
QY 403 GACGCAAGGAGATGCTCATTAAGCAAGATACCTAGAACCAATGATATATTTGAA 462
DB 226 AspSerLysGlyLysCysSerSerLysAlaGluTyLeuArgAspAsnIleMetHisHis 245
QY 463 GCGTTTTCAGGATCGCGGAGAAAACAAGTACTTCTAAAACCACTCAAAATTCACACG 522
DB 246 AlaTyHisAspAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
QY 523 CCCGAATCTAGGCGATGCGCACAGCACTAATGAGACG-----TATACCGTGTGGGATCA 576
DB 266 ProGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyValGlyTrp---Met 284
QY 577 CCATGGATATATCGAACGGACCGTCCGTCATCTATAGTAGAGAAATGATGCCCGC 636
DB 285 ProTrpArgHisTyThrSerThrSerValAsnCysIleValGluGluValGluAlaArg 304
QY 637 TCTGTGTTTCGTTTATCATATTTTCAATGGCC 669
DB 305 SerValTyProTyAspSerPheAlaLeuSer 315
RESULT 9
VGLB_HSVB4
ID VGLB_HSVB4 STANDARD; PRT; 919 AA.
AC P17472;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
OS type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89125704; PubMed=2915378;
RA Riggio M.P., Cullinane A.A., Onions D.E.;

QY	424	TCTAAAGCAGATACCTTTAGAACAATGTTATATGTTGAAGCGTTTGCACAGGGATCGGGA	483
Db	174	SerLysAlaGluTrpLeuArgaspIleMetHisIaIaTrpHisAlaTrpHisaspGluasp	193
QY	484	GRAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCGGAATCTAGGCGATGGCAC	543
Db	194	GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaIaIaTrpGln	213
QY	544	AGCACTAATGAGACG-----TATCCCGTGTGGGATCACCATGGATATATCGAAGCGGA	597

```

598 ACCTCGTCAATTGTATAGTAGAGAAATGGATCCGCTGTGTTCGCTATTCATAT 657
      |||||
233 ThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProIyrAspSer 252
      |||||
568 TTTTCAATGCC 669
      |||||
253 PheAlaLeuSer 256
      |||||

RESULT 10
VGLB_HSV1
ID VGLB_HSV1 STANDARD; PRT; 933 AA.
AC Q0463;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycoprotein B precursor.
GN UL27.
OS Herpesvirus ateles (type 1 / strain Lennette).
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=35243;
ON 111

```

```

SEQUENCE FROM N.A.
MEDLINE-93228440; PubMed-8385913;
Eberle R., Black D.;
"sequence analysis of herpes simplex virus gB gene homologs of two
platyrrhine monkey alpha-herpesviruses.";
Arch. Virol. 129:167-182(1993).
-1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
REQUIRED FOR VIRAL GROWTH.
-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-30 IS THE INITIATOR.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M95785; AAA43839.1; -.
InterPro; IPR000234; Glycoprot_B.
Pfam; PF00606; Glycoprotein_B; 1.
ProDom; PD000693; Glycoprot_B; 1.
Glycoprotein; Transmembrane; Signal.
SIGNAL 1 31
CHAIN 1 31 POTENTIAL.
DOMAIN 32 933 GLYCOPROTEIN B.
DOMAIN 32 753 EXTRACELLULAR (POTENTIAL).
TRANSMEM 754 769 POTENTIAL.
TRANSMEM 775 795 POTENTIAL.
TRANSMEM 798 818 POTENTIAL.
DOMAIN 819 933 CYTOPLASMIC (POTENTIAL).
CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```
FT CARBOHYD 747 747 N-LINKED (GLCNAC.. ) (POTENTIAL).
SQ SEQUENCE 933 AA; 101492 MW; 13A6CF88BBSA18DF CRC64;

Alignment Scores:
Pred. No.: 2,82e-45 Length: 933
Score: 529.50 Matches: 101
Percent Similarity: 62.62% Conservative: 28
Best Local Similarity: 49.03% Mismatches: 74
Query Match: 44.12% Indels: 3
DB: 1 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSV1 (1-933)
QY 52 GGTAGCAATCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTC 111
Dy 108 GlyThrArgSerAlaAlaArgGln-----LeuArgGluSerLeuArgArgile 124
QY 112 CAGTTGTCTGAGGAGACTACGTTTATCTTGTGCCCCCACCAGTGGTTCACCGCG 171
Dy 125 GlnAlaGluTyrAlaAlaSerAlaPheTyrValCysProProThrGlyAlaThrVal 144
QY 172 ATCCGCTCAGAACCCCGCGAAATGTCCGAACTAGAAAAGCCAGCGTGGGTGAA 231
Dy 145 ValGlnPheGluProArgProCysProAspValAlaAlaGlyLysAsnPhetrGlu 164
QY 232 GGAATCGGATATTATTAAAGAAATATCATGTCATATAAATTTAAAGTACGCTTAT 291
Dy 165 GlyIleAlaValPheLysGluAsnIleAlaProTyrLysPheThrAlaThrLysTyr 184
QY 292 TATAAATATCATTCACAGCAGCATGGAGGGGCGACATATAGACATCACTAAT 351
Dy 185 TyrLysGluIleThrValSerGlnThrTrpGlnGlySerArgTyrLeuGlnLeuThrGly 204
QY 352 CGATACAGATAGGACCGCGTTCATTTCCATTAAGAGATACAGGATCTAATCGACGCAA 411
Dy 205 LeuTyrAsnAspArgAlaProValProPheSerGluIleThrAspLeuIleasnGlyLys 224
QY 412 GGAATGCTCATCTAAAGCAGATACCTTGAACAATGTATATGTGAAGCGTTTGC 471
Dy 225 GlyArgCysArgSerAspValThrTyrThrArgSerGlnArgValThrAlaTyrAsp 244
QY 472 AGGGATCGGGAGAAAACAAGTACTCTTAAACCATCAAAATTCACACGCCCGAATCT 531
Dy 245 GlyAspGluTrpGlyArgGluValAlaLeuValProAlaLysThrSerThrProAsnSer 264
QY 532 AGGGATCGGCACACGACTAATGAGCGTATACCGGTGTGGGGATCACCATGGATATATCGA 591
Dy 265 ArgGlyTyrTyrThrThrAspArgValTyrAlaProAsnAlaHisAlaGlyPheTyrLys 284
QY 592 ACGGGAACCTCGTCAATGTATAGTAGGAATGGATGCCGCTGTGTTCGCTAT 651
Dy 285 ThrGlyThrThrValAsnCysIleValGluGluMetGluAlaArgSerAlaPheProTyr 304
QY 652 TCATATTTCGAATGGCC 669
Dy 305 AspSerPheValLeuAla 310

RESULT 11
VGLB_HSVSM
ID VGLB_HSVSM STANDARD; PRT; 920 AA.
AC Q04464;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN UL27.
OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset
OC Herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10353;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE-93228440; PubMed-8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gB gene homologs of two
RL Arch. Virol. 129:167-182(1993).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59786; AAA43841.1; -
DR InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT CHAIN 1 28
FT DOMAIN 29 920 GLYCOPROTEIN B.
FT TRANSMEM 740 755 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 761 781 POTENTIAL.
FT TRANSMEM 784 804 POTENTIAL.
FT DOMAIN 805 920 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 583 583 N-LINKED (GLCNAC.. ) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC.. ) (POTENTIAL).
SQ SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;

Alignment Scores:
Pred. No.: 6,34e-45 Length: 920
Score: 526.00 Matches: 96
Percent Similarity: 65.28% Conservative: 30
Best Local Similarity: 49.74% Mismatches: 67
Query Match: 43.83% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVSM (1-920)
QY 91 AGAGAAGTGTTCGAGCGTCCAGTTGTCTGAGGAGAGTCTPACGTTTATCTTTGTC 150
Dy 109 ArgGluSerValArgGlnIleArgAlaGluAsnAlaThrSerMetPheTyrValCysPro 128
QY 151 CCACGAGTGGTTCACCGTGATCGCTGTAGAACCGCGGAAATGTCCGAACTAGA 210
Dy 129 ProProThrGlyAlaThrValValGlnPheGluGluProArgProCysProAspValAla 148
QY 211 AAAGCCACCGAGTGGGTGAAGGAATCGGATATTTTAAAGAGAATATCATGTCATAT 270
Dy 149 AlaGlyLysAsnPhetrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyr 168
QY 271 AAATTTAAAGTGACGCTTTATATATAAAATATCATTCAGACGACATGGACGGGAGC 330
Dy 169 LysPheThrAlaThrMetTyrTyrLysGluIleThrValThrGlnThrTrpGlnGlySer 188
QY 331 ACATATAGACATCACTAATCGATATACATAGACGCGCGCTTTCCATTGAAAGATC 390
Dy 189 ArgTyrLeuGlnLeuThrGlyLeuTyrAsnAspArgAlaProValProPheGluGluIle 208
QY 391 ACGGATCTAATCGCGGCAAGGAGATGCTCATCTAAGCAGAGATACCTTGAACAAT 450
Dy 209 ThrAspValIleAsnAlaLysGlyLeuCysArgSerAspValThrTyrValArgSerGln 228
```

QY	451	GTATATGTTGAAGCGTTTGACAGCGATGCGGAGAAACAACTACTTCTAAACCATCA	510
Db	229	Arg-gValThrAlaIaTyAspArgSpGLuTrpGlyArgGluValLysLeuValProSer	248
QY	511	AAATTCACACGCCCGCAATCTAGGCGCATGCGACACGACTAATCAGACGCTATACCGGTGG	570
Db	249	LysThrSerThrProAsnSerArgGlyTrpTyThrThrAspArgMetTyAlaProAsn	268
QY	571	GGATCACCATGGATATATCGAACCGGGAACCTCCGCTCAATTGTATAGTAGAGAAATGGAT	630
Db	269	AlaHisAlaGlyPheTyAlaGlyThrThrValAsnCysIleValGluGluValGlu	288
QY	631	GCCCGCTGCTGTTTCGTTATTCATATATTGCAATGGCC	669
Db	289	AlaArgSerAlaTyProTySerAsnPheValLeuAla	301
RESULT 12			
VGLB_HSV1			
ID	VGLB_HSV1	STANDARD;	PRF; 980 AA.
AC	PI8050;		
DT	01-NOV-1990 (Rel. 16, Created)		
DT	01-NOV-1990 (Rel. 16, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Glycoprotein B precursor (glycoprotein 14).		
DE	GB OR GP14 OR 33.		
OS	Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).		
OS	Viruses: dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphan herpesvirinae; Varicellovirus.		
ON	NCBI_Taxid=10327;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=89279217; PubMed=2543744;		
RA	Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,		
RA	Woodworth L.M.;		
RT	"Identification and nucleotide sequence of a gene in equine		
RT	herpesvirus 1 analogous to the herpes simplex virus gene encoding the		
RT	major envelope glycoprotein gb.;"		
RL	J. Gen. Virol. 70:383-394(1989).		
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement).		
CC	or send an email to license@isb-sib.ch .		
CC	-----		

DR	EWBL:	D00401;	BAA00304.1;	ALT_SEQ.
DR	PIR:	A31241;	VGBEZH.	
DR	InterPro:	IPR002334;	Glycoprot_B.	
DR	Pfam:	PF00606;	Glycoprotein_B; 1.	
DR	ProDom:	PD000693;	Glycoprot_B; 1.	
KW			Glycoprotein; Transmembrane; Signal.	
FT	SIGNAL	1	POTENTIAL.	
FT	CHAIN	87	980	GLYCOPROTEIN B.
FT	DOMAIN	87	852	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	853	870	POTENTIAL.
FT	DOMAIN	871	980	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	165	165	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	275	275	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	380	380	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	423	423	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	497	497	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	514	514	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	560	560	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	727	727	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	749	749	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	980 AA;	109931 MW;	94B69AF882FB6AC2 CRC64:

Alignment Scores:

Pred. No.:	7.24e-45	Length:	980
Score:	525.50	Matches:	109
Percent Similarity:	57.94%	Conservative:	37
Best Local Similarity:	43.25%	Mismatches:	71
Query Match:	43.79%	Indels:	35
DB:	1	Gaps:	6

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSVB1 (1-980)			
QY	13	AGGCGGAATTCATATTTTCTTATAGTATTCTATATGGT-----	54
DB	66	ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla	85
QY	55	-----ACGAATCATCTCCGAGTACCCAAAATGTGACATCA-----	90
DB	86	ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis	105
QY	91	-----AGAGAAGTGTGTT	102
DB	106	GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal	125
QY	103	TCGAGCGTCCAGTTGCTGAGAGAGAGCTACGTTTATCTTTGTCCTCCACCATGGGT	162
DB	126	ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly	145
QY	163	TCAACCGTATCCGTCTAGAACCCGCGGAAATGTCGCCAGACCTAGAAAACCCACCGAG	222
DB	146	SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn	165
QY	223	TGGGGTGAAGAAATCGCGATATTATTAAAGAGAATATCAGTCCATATAAATTTAAAGTG	282
DB	166	PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla	185
QY	283	ACGCTTTATTAAATATATCATTCAGACGAGCATCGACGGGACCATATATAGACAG	342
DB	186	AsnValTyrTyrLysAspIleValThrArgValTyrLysGlyTyrSerHisThrSer	205
QY	343	ATCACTAATCGATATACAGATAGGACGCCGCTTCCATTTGAAGAGATCAGGATCTAATC	402
DB	206	LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuIle	225
QY	403	GACGGCAAGGAAGATGCTCATCTAAACAGATACCTTGAACAACATGTATATGTTGAA	462
DB	226	AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMethHis	245
QY	463	CGCTTTGACAGGGATCGGGAGAAAACAAGTA---CTTCTAAACCATCAAAATTCAC	519
DB	246	AlaTyrHisAspAspGluAspGluValGluLeuAspLeuCysArgProSer---LeuGln	264
QY	520	ACGCCCCGAATCTAGGGCATGCACACACTAATGAGAGC-----TATACCGTGTGGGA	573
DB	265	LeuArgGlyAlaArgAlaTprGluThrThrAsnAspThrThrSerTyrValGlyTyr---	283
QY	574	TCACCATGATATATCGAAGCGGAAACCTCCGCTCAATTTGTATAGTAGAGAAATGATGCC	633
DB	284	MetProTprArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAla	303
QY	634	CGCTCTCGTTTCCGTATTCATATTTGCCATGGCC	669
DB	304	ArgSerValTyrProTyrAspSerPheAlaLeuSer	315

RESULT 13
VGLB_HSV2S
ID VGLB_HSV2S
AC P24994; STANDARD; PRT; 885 AA.
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
GN Glycoprotein B precursor.
DE GB OR UL27.
OS Herpes simplex virus (type 2 / strain S48) (Simlan agent 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.


```

OX NCBI_TaxID=10316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B264;
RX MEDLINE=91374035; PubMed=1895066;
RA Borchers K., Weigelt W., Buhk H.-J., Ludwig H., Mankertz J.;
RT "Conserved domains of glycoprotein B (gB) of the monkey virus, simian
RT agent 8, identified by comparison with herpesvirus gBs.";
RL J. Gen. Virol. 72:2299-2304(1991).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GG, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56935; CAA40256.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 34
FT CHAIN 35 885
FT DOMAIN 35 715
FT TRANSMEM 716 731
FT TRANSMEM 737 756
FT TRANSMEM 760 780
FT DOMAIN 781 885
FT CARBOHYD 68 68
FT CARBOHYD 122 122
FT CARBOHYD 379 379
FT CARBOHYD 411 411
FT CARBOHYD 659 659
FT CARBOHYD 885 885
SQ SEQUENCE 885 AA; 97811 MW; 39B495B329AB94E4 CRC64;

Alignment Scores:
Pred. No.: 1.04e-42 Length: 885
Score: 504.00 Matches: 94
Percent Similarity: 62.75% Conservative: 34
Best Local Similarity: 46.08% Mismatches: 76
Query Match: 42.00% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSV2S (1-885)
QY 58 AACTCATCTCCGAGTACCAAAATGTGACATCAAGAGAGTGTGTTTCGAGCGTCAGTTG 117
Db 68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleVal 87
QY 118 TGTGAGGAAGACTACGTTTATCTTTGTCGCCACCGAGTGGTTCACCGTCGCGT 177
Db 88 ArgAspGlyAlaThrPheTyrValCysProProThrGlyAlaThrValGln 107
QY 178 CTAGAACCCCGCGAAATGTGCCCCAACCTAGAAAGCCAGCGAGTGGGTGAAGATC 237
Db 108 PheGluGlnProArgProCysProArgAlaProAspGlyGlnAsnTyrThrGluGlyIle 127
QY 238 GCGATATTATTAAAGAAATACAGTCATATAATTTAAAGTCACGCTTTATATATAA 297
Db 128 AlaValPheValGluAlaProTyrValPheValGluAlaThrMetTyrTyrIys 147
QY 298 AATATCATTACGACGACGATGGAGCGGCGGACATATAGACATCAATCGATAT 357
Db 148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetClyIlePhe 167

```

```

QY 358 ACAGATAGACGCCCGTTTCCATTTCAAGAGATCAGCGATCAATCGACGGCAAGAAGA 417
Db 168 GluAspArgAlaProValProPheGluGluValMetAspLysIleAsnAlaLysGlyVal 187
QY 418 TGTCTCATCTAAAGCAAGATACCTTGAACAATGATATATGTTGAAGCGTTTGACAGGAT 477
Db 198 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHisArgAsp 207
QY 478 GCGGGGAGAAACAAGTACTTCTTAAACCATCAAAATTCACACACCCCGAATCTAGGCGA 537
Db 208 AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly 227
QY 538 TGGCACACGACGTAATGACGACGATATACCGTGTGGGATACCATGATATATCGAACGGGA 597
Db 228 TrpHisThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly 247
QY 598 ACCTCCGTCATTTGATAGTAGAGAAATGGATGCCGTCGTCTTCCGTCATTCATAT 657
Db 248 ThrThrValAsnCysIleValGluGluAlaArgSerValTyrProTyrAspGlu 267
QY 658 TTGTCATGCGC 669
Db 268 PheValLeuAla 271

RESULT 14
VGLB_HSV1F
ID VGLB_HSV1F STANDARD; PRT; 903 AA.
AC P06436;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain F).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85083254; PubMed=2981343;
RA Pellett P.E., Kousoulas K.G., Pereira L., Roizman B.;
RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
RT primary sequence and predicted protein structure of the wild type and
RT of monoclonal antibody-resistant mutants.";
RL J. Virol. 53:243-253(1985).
RN [2]
RP SEQUENCE OF 1-176 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14164; AAA45776.1;
DR EMBL: M21633; AAA45788.1;
DR PIR: A03750; VGBEB1.
DR InterPro; IPR000234; Glycoprot_B.

```

DR	Pfam:	PF00606;	Glycoprotein_B; 1.
DR	ProDom:	PD00693;	Glycoprot_B; 1.
KW	Glycoprotein;	Transmembrane;	Signal.
FT	SIGNAL	1	29
FT	CHAIN	30	903
FT	DOMAIN	31	729
FT	TRANSMEM	730	745
FT	TRANSMEM	751	770
FT	TRANSMEM	774	794
FT	DOMAIN	795	903
FT	CARBOHYD	86	86
FT	CARBOHYD	140	140
FT	CARBOHYD	397	397
FT	CARBOHYD	429	429
FT	CARBOHYD	488	488
FT	CARBOHYD	673	673
SQ	SEQUENCE	903 AA;	100104 MW; 73BDCAT813DB35E8 CRC64;

Alignment Scores:			
Pred. No.:	1.05e-42	Length:	903
Score:	504.00	Matches:	93
Percent Similarity:	63.11%	Conservative:	37
Best Local Similarity:	45.15%	Mismatches:	76
Query Match:	42.00%	Indels:	0
DB:	1	Gaps:	0

US-09-147-052-3_COPY_1_672 (1-672) x VGLB_HSV1F (1-903)

QY	52	GTGACGAAGTCTCCGAGTACCNAATGTGACATCAGACAAGTTGTTTCGACGGTC	111
DB	84	GlyAspAsnAlaThrValAlaAlaGlyHisAlaThrLeuArgGluHisLeuArgAspIle	103
QY	112	CAGTTGTCTGAGGAAGTCTACGTTTATCTTTGTCCCCCACACAGTGCGGTTCACCCGNG	171
DB	104	LysAlaGluAsnThrAspAlaAsnPheTyValCysProProThrGlyAlaThrVal	123
QY	172	ATCCGCTAGAACCCGCCGAAATGTCCCGAAGCTAGAAAACCCACCGAGTGGGTGAA	231
DB	124	ValGlnPheGluInProArgArgCysProThrArgProGluGlyGlnAsnTyThrGlu	143
QY	232	GAATCGCGATATTATTAAGACGAATATCAGTCCATATAAATTTAAAGTAGCGCTTAT	291
DB	144	GlyIleAlaValPheLysGluAsnIleAlaProTyrlLysPheLysAlaThrMetTy	163
QY	292	TATAAAATATCATTCAGACGACATGGACGGGACGACATATAGACAGATCACTAAT	351
DB	164	TyrLysAspValThrValSerGlnValTrpPheGlyHisArgTySerGlnPheMetGly	183
QY	352	CGATATACAGATAGACGCCGCTTTCATTAAGAGATACCGGATCTAATCAGCGCAA	411
DB	184	IlePheGluAspArgAlaProValProPheGluValIleAspLysIleAsnAlaLys	203
QY	412	GGAAGTCTCATCTMAAGCAAGATACCTTAGAACCAATGTATATGTTGAAGCGTTGAC	471
DB	204	GlyValCysArgSerThrAlaLysTyValArgAsnAsnLeuGluThrThrAlaPheHis	223
QY	472	AGGGATCGGGAGAAAAAAGTACTTCTAAAACCATCAAAATTCACACGCCCGAATCT	531
DB	224	ArgAspAspHisGluThrAspMetGluLeuLysProAlaAsnAlaAlaThrArgThrSer	243
QY	532	AGGGCATGGCACAGCTAATAGACGCTATACCGTGTGGGGATCCACCATGATATATCA	591
DB	244	ArgGlyTyPHisThrThrAspLeuLysTyAsnProSerArgValGluAlaPheHisArg	263
QY	592	ACGGGAACCTCCGTCANTGTATAGTAGAGAAATGGATCGCGCTCTGTGTTCCGAT	651
DB	264	TyrGlyThrThrValAsnCysIleValIsGluValAspAlaArgSerValTyProTy	283
QY	652	TCATATTTTGCATGGCC	669
DB	284	AspGluPheValLeuAla	289

RESULT 15

```
QY 172 ATCCGCTAGACCCGCCGAAATGTCCTCCGACCTAGAAACCCACCCAGTGGGTGAA 231
Db :::: ||| |||||:::||||| ::: |||
125 ValGlnPheGluGlnProArgCysProThrArgProGluGlnAsnTyrThrGlu 144
QY 232 GGAATCGCGATATTATTAAGAGAAATATCAGTCCATATAAATTTAAAGTGACGCTTTAT 291
Db |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
145 GlyIleAlaValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyr 164
QY 292 TATAAAATATCATTCAGACGACATGGACGGGACGACATATAGACAGATCACTAAT 351
Db |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
165 TyrLysAspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGly 184
QY 352 CGATATACAGATAGGACGCCGCTTCCATTGAAGATCACGGATCTAATCGACGCCAAA 411
Db ::: ||||| ||||| |||||::: ||| |||||::: ||| |||||::: |||
185 IlePheGluaspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaLys 204
QY 412 GGAAGATGCTCATCTAAACAAAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTGAC 471
Db ||| ||| ||| |||||:::||||| |||||::: ||||| |||||::: |||||
205 GlyValCysArgSerThrAlaLysTyrValArgAsnAsnLeuGluThrThrAlaPheHis 224
QY 472 AGGGATCGCGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCT 531
Db ||||| ||| ::: |||||::: |||||::: ||| |||||::: |||
225 ArgAspHisGluThrAspMetGluLeuLysProAlaAsnAlaAlaThrArgThrSer 244
QY 532 AGGGCATGCCACACGACTAATGAGCGTATACGGTGTGGGGATCACCATGGATATATCGA 591
Db ||| |||||::: ||| |||||::: ||| |||||::: ||| ::: |||
245 ArgGlyTrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArg 264
QY 592 ACGGGAACCTCGCTCAATGTATAGTAGAGAAATGGATCCCGCTCTGTTCGGTAT 651
Db |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
265 TyrGlyThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyr 284
QY 652 TCATATTTTGCATGGCC 669
Db ||| ::: |||
285 AspGluPheValLeuAla 290
```

Search completed: October 8, 2003, 17:56:53
Job time : 19.8306 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:32 ; Search time 22.1621 Seconds
(without alignments)
5832.060 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200

Sequence: 1 atgcactatttttagcgga.....catattttgcaatggccaat 672

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p_model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154341_29860/app_query.fasta_1.5980
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_287 -runat_08102003_154341_29860 -NCPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	100.0	865	1	VBERSB
2	667	55.6	948	2	A56602
3	606	50.5	618	1	VBEB31
4	605	50.4	913	1	VBEBPS
5	583	48.6	928	1	VBEBEG
6	583	48.6	943	2	B48474
7	564.5	47.0	932	1	VBEBEC
8	563.5	47.0	979	2	JH0109
9	563.5	47.0	980	1	VBEBE6
10	559.5	46.6	919	1	VBEBEQ
11	559.5	46.6	975	2	T42576
12	529.5	44.1	933	1	B48349
13	525.5	43.8	980	1	VBEB2H
14	524	43.7	920	1	D48349

15	504	42.0	885	1	VBESA	glycoprotein B pre
16	504	42.0	903	1	VBEB1	glycoprotein B pre
17	504	42.0	903	1	VBEB1	glycoprotein B pre
18	504	42.0	904	1	VBEBW7	glycoprotein B pre
19	500	41.7	904	1	VBEBK2	glycoprotein B pre
20	500	41.7	904	1	VBEBB2	glycoprotein B pre
21	499	41.6	917	1	VBEBH	glycoprotein B pre
22	426.5	35.5	883	1	VBEBIS	glycoprotein B pre
23	422	35.2	883	1	VBEBIL	glycoprotein B pre
24	420	35.0	873	2	S26690	glycoprotein B - i
25	219	18.2	874	2	S25530	glycoprotein B - b
26	217.5	18.1	874	2	S55602	glycoprotein B - e
27	211.5	17.6	808	1	VBESM	glycoprotein B pre
28	211.5	17.6	830	1	B44047	glycoprotein B pre
29	211.5	17.6	830	2	T43999	glycoprotein B [im
30	211.5	17.6	830	2	T44186	probable glycoprot
31	209.5	17.5	822	2	T41941	glycoprotein B - h
32	209.5	17.5	830	1	A44047	glycoprotein B pre
33	204	17.0	854	2	T03107	glycoprotein B - a
34	202.5	16.9	907	1	VBETE	glycoprotein B pre
35	201.5	16.8	831	1	VBEB6S	glycoprotein B - h
36	201.5	16.8	857	1	QBEB1L	glycoprotein B - h
37	196.5	16.4	807	2	T42924	glycoprotein B - a
38	194.5	16.2	906	1	VBEB1	glycoprotein B pre
39	180	15.0	928	1	VBEMC	glycoprotein B pre
40	89	7.4	326	2	AF2082	iron(III) dictitrat
41	89	7.4	848	1	VBEBD3	glycoprotein H pre
42	88.5	7.4	1163	2	A36685	205K microtubule-a
43	88	7.3	416	1	A31959	lysosome-associate
44	86.5	7.2	385	2	A40359	aliphatic amidase
45	86.5	7.2	1008	2	T18727	hypothetical prote

ALIGNMENTS

RESULT 1

VBERSB
glycoprotein B precursor - Marek's disease virus (strain RB1B)
C:Species: Marek's disease virus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A32402; B32402
R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.
J. Gen. Virol. 70, 1789-1804, 1989
A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog
A:Reference number: A32402; MUID:89293086; PMID:2544666
A:Accession: A32402
A:Molecule type: DNA
A:Residues: 1-865 <ROS>
A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837
A:Accession: B32402
A:Molecule type: protein
A:Residues: 250-271;304-330 <ROS2>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-2l/Domain: signal sequence #status predicted <SIG>
F:22-865/Product: glycoprotein B #status predicted <GPB>
F:709-728/Domain: transmembrane #status predicted <TN1>
F:732-752/Domain: transmembrane #status predicted <TN2>
F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Alignment Scores:			
Pred. No.:	7.57e-112	Length:	865
Score:	1200.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-147-052-3_COPY_1_672 (1-672) x VBERSB (1-865)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTATAGTATCTATGTTAGGACGAC 60

DB 1 MethistyrPheArgAsnCysIlePheLeuValIleLeuTyrGlyThrAsn 20

QY 61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTGTCTTCGAGCGTCCAGTGTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTACGTTTATCTTGTCCCCACCAAGTGGGTCAACCGTGATCCGCTCA 180
 Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAACCCCGCGAAATGTCCCAACCTAGAAAAGCCAGGAGTGGGTGAAGAAATCGCG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
 QY 241 ATATTATTAAAGAGATATACGTCCATATAATTTAAAGTGACGCTTTATTTATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyLysPheLysValThrLeuTyTyLysAsn 100
 QY 301 ATCATTCAGACGACGATGGGAGCGGAGACATATAGACAGATCACTAATCCATATACA 360
 Db 101 IleIleGlnThrThrThrThrThrGlyThrThrThrThrThrThrThrThrThrThr 120
 QY 361 GATAGGACGCCGTTTCCATTAAGAGATCACGGATCTAATCGACGGCAAGAGATGC 420
 Db 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
 QY 421 TCATCTAAACGAGATACCTTAGAAACATGTATATGTGAACGGTTTGACAGGGATGG 480
 Db 141 SerSerLysAlaArgTyLeuArgAsnValTyValGluAlaPheAspArgAspAla 160
 QY 481 GGAGAAAACAACTACTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGCGATGG 540
 Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
 QY 541 CACACGACTAATGAGACGATACCGTGTGGGATACCATGATGATATATGAAAGCGGAAC 600
 Db 181 HisThrThrAsnGluThrThrValTrpGlySerProTrpIleTyArgThrGlyThr 200
 QY 601 TCGCTCAATTTGTATAGTACGAGAAATGGATGCCGCTCTGCTTCCGTATTCATATTT 660
 Db 201 SerValAsnLysIleValGluGluMetAspAlaArgSerValPheProTySerTyPhe 220
 QY 661 GCAATGGCCCAAT 672
 Db 221 AlaMetAlaAsn 224

RESULT 2

A56602
 glycoprotein B homolog precursor - feline herpesvirus 1
 C:Species: feline herpesvirus 1
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999
 C:Accession: A56602
 R:Naeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Niihara, M.; Kai, Arch. Virol. 127, 387-397, 1992
 A:Title: Identification and nucleotide sequence of a gene in feline herpesvirus type 1 H A:Reference number: A56602; MUID:93090104; PMID:1333759
 A:Accession: A56602
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-948 <NAE>
 A:Cross-references: GB:S49775; NID:Q261094; PIDN:AAB24381.1; PID:G261095
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:120002, NCBIP:120003)
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein

Alignment Scores:

Pred. No.:	2,08e-58	Length:	948
Score:	667.00	Matches:	128
Percent Similarity:	59.43%	Conservative:	39
Best Local Similarity:	45.55%	Mismatches:	54
Query Match:	55.58%	Indels:	60
DB:	2	Gaps:	3

RESULT 3

VGBE31

glycoprotein B - human herpesvirus 3
 N:Alternate names: glycoprotein II
 C:Species: human herpesvirus 3, varicella-zoster virus
 C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: E27214
 R:Davidson, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657; PMID:3018124

US-09-147-052-3_COPY_1_672 (1-672) x A56602 (1-948)

QY 7 TATTTAGGCGGAATTCATATTT----- 30
 Db 22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
 QY 31 -----TTCCTTATAGTT 42
 Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyValSerPheIleTrpIle 61
 QY 43 ATTCTATAT----- 51
 Db 62 ValLeuPheLeuValGlyProArgProValGluGlyClnSerGlySerThrSerGluGln 81
 QY 52'-----GGTACGAACCTCATCTCCGACTACCCAA 78
 Db 82 ProArgArgThrValAlaAlaThrProGluValGlyGlyThrProProLysProThrThrAsp 101
 QY 79 AATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTCTCTCGAGGAAG----- 129
 Db 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
 QY 130 ---TCTACGTTTATCTTGTCCCCACCAAGTGGTTCACCGTGATCCGCTCTAGAACCG 186
 Db 122 ProSerThrPheTyLeuMetCysProProSerGlySerThrValValArgLeuGluPro 141
 QY 187 CCGCGAAATGTCCCAACCTAGAAAAGCCAGTGGGGTGAAGGAATCGCATATTA 246
 Db 142 ProArgAlaCysProAspTyLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
 QY 247 TTTAAAGAGATATCATATTAATTTAAAGTACGCTTTATTTATAAAATATCATTT 306
 Db 162 PheLysGluAsnIleAlaProTyLysPheLysAlaAsnIleTyTyLysAsnIleIle 181
 QY 307 CACACACGACATACGAGCGGACGACATATAGACAGATCACTAATCATATACAGATAGG 366
 Db 182 MetThrThrValTrpSerGlySerSerTyAlaValThrThrAsnArgTyThrAspArg 201
 QY 367 AGCCCGTTTCCATTAAGAGATACGATCTAATCGACGGCAAGGAAGATCGTCTATCT 426
 Db 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeuSer 221
 QY 427 AAAGCAAGATACCTTAGAACATGTATGTGTGAAGCGTTTCACAGGATCGGGAGAA 486
 Db 222 LysAlaAspTyValArgAsnAsnTyGlnPheThrAlaPheAspArgAspGluAspPro 241
 QY 487 AAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGCGATGGCACAG 546
 Db 242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTrpHisThr 261
 QY 547 ACTAATGAGACGTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACCTCCGTC 606
 Db 262 ThrAsnGluThrTyThrLysIleGlyAlaAlaGlyPheHisHisSerGlyThrSerVal 281
 QY 607 AATTGTATAGTAGAAGAAATGGATGCCGCTCTGTTTCCGTATTCATATTTTGAATG 666
 Db 282 AsnCysIleValGluGluValAspAlaArgSerValTyProTyArgSerPheAlaIle 301
 QY 667 GCC 669
 Db 302 Ser 302

A:Accession: E27214

A:Molecule type: DNA

A:Residues: 1-868 <DAY>

A:Cross-references: EMBL:X04370; NID:959989; PIDN:CAA27914.1; PID:g60020

C:Genetics:

C:Gene: 31

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:697-713/Domain: transmembrane #status predicted <TM1>

F:726-742/Domain: transmembrane #status predicted <TM2>

F:84,194,372,416,440,494,557,623,781,836,864/Binding site: carbohydrate (Asn) (covalent)

Alignment Scores:

Pred. No.:	2,676-52	Length:	868
Score:	606.00	Matches:	107
Percent Similarity:	74.09%	Conservative:	36
Best Local Similarity:	55.44%	Mismatches:	50
Query Match:	50.50%	Indels:	0
DB:	1	Gaps:	0

US-09-147-052-3_COPY_1_672 (1-672) x VGBE31 (1-868)

QY	91	AGAGAAGTGTGTTGAGCGTCCAGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTC	150
DB	41	ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro	60
QY	151	CCACGAGTGGTTCAACCGTGATCGCTGTAGAACCGCGCGAATAATGTCGGAACCTAGA	210
DB	61	ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis	80
QY	211	AAAGCCACGAGTGGGTGAGGATCGCGATATATTAAAGAGATATCAGTCCATAT	270
DB	81	LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaAlaTyr	100
QY	271	AAATTAAAGTACGCTTTATATAAAATATCATTCAGACGACACATGGACGGGACG	330
DB	101	LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTipAlaGlySer	120
QY	331	ACATATACAGATACATTAATCGATATACATAGGACCGCGGTTTCCATTGAGAGATC	390
DB	121	SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle	140
QY	391	ACGGATCAATACGACGGCAAGGAGATGCTCATCTAAAGCAAGATACCTTAGAACCAAT	450
DB	141	ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn	160
QY	451	GTATATGTGAAGCGTTTGACAGGATCGCGGAGAAAACAAAGTACTTCTAAACCATCA	510
DB	161	HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer	180
QY	511	AAATTCAACACCGCGAATCTAGGCGATGGACAGCATATGAGACGTATACCGTGTGG	570
DB	181	LysTyrAsnSerValGlySerLysAlaThrHisThrAsnAspThrTyrMetValAla	200
QY	571	GGATCACCATGATATACGACGGAACCTCCGTCATTTAGTAGAGAAATGAT	630
DB	201	GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluValGlu	220
QY	631	GCCCCCTGTGTTTCCGTTATCATATTTCGATGGCC	669
DB	221	AlaArgSerIlePheProTyrAspSerPheGlyLeuSer	233

RESULT 4

VGBE31

glycoprotein gII precursor - suid herpesvirus 1

C:Species: suid herpesvirus 1

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A29159

R:Robbins, A.K.; Dorney, D.J.; Wathen, M.W.; Whealy, M.E.; Gold, C.; Watson, R.J.; Holla

J. Virol. 61, 2691-2701, 1987

A:Title: The pseudorabies virus gII gene is closely related to the gB glycoprotein gene

A:Reference number: A29159; MUID:87284141; PMID:3039163

A:Accession: A29159

A:Molecule type: DNA

A:Residues: 1-913 <ROB>

A:Cross-references: GB:M17321; NID:g334053; PIDN:AAA7465.1; PID:g334054

A:Note: the authors translated the codon GAC for residue 860 as Asn

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-40/Domain: signal sequence #status predicted <SIG>

F:751-819/Domain: transmembrane #status predicted <TM1>

F:820-913/Domain: intracellular #status predicted <INT>

F:151,261,441,516,633,697/Binding site: carbohydrate (Asn) (covalent) #status predict

Alignment Scores:

Pred. No.:	3,396-52	Length:	913
Score:	605.00	Matches:	112
Percent Similarity:	72.12%	Conservative:	38
Best Local Similarity:	53.85%	Mismatches:	54
Query Match:	50.42%	Indels:	4
DB:	1	Gaps:	2

US-09-147-052-3_COPY_1_672 (1-672) x VGBE31 (1-913)

QY	52	GGTACGAACACTCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAAGTGTGTCG	105
DB	95	GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg	114
QY	106	ACGCTCCAGTTGTCGAGGAAGAGTCTACGTTTATCTTTTGTCCCCACCAGTGGTTCA	165
DB	115	AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer	132
QY	166	ACGCTGATCCGCTCTAGAACCCCGCGAAATGTCGCCAACCTAGAAAACCCACGAGTGG	225
DB	133	ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGlnGlyArgAsnPhe	152
QY	226	GGTGAAGGAATCGGATATATTAAAGAGATATCATCTCCATATAAATTTAAAGTACG	285
DB	153	ThrGluGlyIleAlaValLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis	172
QY	286	CTTTATTAATAATATCATTCAGACGACGATGGCGGAGCAGACATATAGACAGATC	345
DB	173	IleTyrTyrLysAsnValIleValThrThrValTrpSerGlySerThrTyrAlaAlaIle	192
QY	346	ACTAATCATATACAGATAGGACCGCGTTTCCATTGAAGAGATCAGCGATCTAATCCAC	405
DB	193	ThrAsnArgPheThrAspArgValProValProValGluGluIleThrAspValIleAsp	212
QY	406	GGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACAATGTATATGTTGAAGCG	465
DB	213	ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla	232
QY	466	TTTGACAGGATGCGGGAGAAAACAAAGTACTTCTAAACCATCAAAATTCACACGCC	525
DB	233	PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu	252
QY	526	GAATCTAGGCGATGCACACGACTAATGAGCGTATCCGTCGTGGGATCACCATTGATA	585
DB	253	GlyThrArgGlyTrpHisThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe	272
QY	586	TATCGAAGCGGAACCTCCGTCATATGATAGAGAAATGGATCCCGCTCTGTGTT	645
DB	273	TyrHisThrGlyThrSerValAsnCysIleValGluValGluAlaArgSerValTyr	292
QY	646	CCGTATTCATATTTTGCATGCC	669
DB	293	ProTyrAspSerPheAlaLeuSer	300

RESULT 5

VGBE31

glycoprotein gI precursor - bovine herpesvirus 1

N:Alternate names: glycoprotein Ila; glycoprotein 16; glycoprotein gi30; glycoprotein

C:Species: bovine herpesvirus 1

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C:Accession: A31166

R:Misra, V.; Nelson, R.; Smith, M.

Virology 166, 542-549, 1988

A:Title: Sequence of a bovine herpesvirus type-1 glycoprotein gene that is homologous to the gene for the glycoprotein gp116 of the pseudorabies virus

A:Reference number: A31166; PMID:89020821; PMID:2845660

A:Accession: A31166

A:Molecule type: DNA

A:Residues: 1-928 <MIS>

A:Cross-references: GB:M23257; NID:g340858; PIDN:AAA6013.1; PID:g511852

A:Note: the authors translated the codon CTG for residue 534 as Ser

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-67/Domain: signal sequence #status predicted <SIG>

F:68-928/Product: glycoprotein g1 #status predicted <GGI>

F:756-824/Domain: transmembrane #status predicted <TMN>

F:105,153,442,484,579,637,703/Binding site: carbohydrate (Asn) (covalent) #status predicted <CAR>

Alignment Scores:

Pred. No.:	5 46e-50	Length:	928
Score:	583.00	Matches:	107
Percent Similarity:	68.45%	Conservative:	34
Best Local Similarity:	51.94%	Mismatches:	65
Query Match:	48.58%	Indels:	0
DB:	1	Gaps:	0

US-09-147-052-3_COPY_1_672 (1-672) x VGBEG (1-928)

QY	52	GGTAGCAACTCATCTCCGAGTACCGCAAATGTGACATCAAGAAGAAGTTGTTCCTGCAGCGTC	111
Db	97	GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaLeuArgLeuAla	116
QY	112	CAGTTGTCTGAGGAGAGCTCTACCTTTTATCTTTTGTCCTCCACCAGTCAGTGGTTCAACCGTG	171
Db	117	GlnAlaAlaGlyIleuAsnSerArgPheValCysProProSerGlyAlaThrVal	136
QY	172	ATCCGCTAGAACCCGCCCGAAAAATGTCCTCCGAACCTAGAAAAGCCACCGAGTGGGGTGAA	231
Db	137	ValArgLeuAlaProAlaArgProCysProGluTyrrGlyLeuGlyArgAsnTyrrThrGlu	156
QY	232	GGATCGCGGATATTATTTAAAGAGAATATCAGTCCATATAAATTTAAAGTAGCGCTTAT	291
Db	157	GlyIleGlyIleTyrrLysGluAsnIleAlaProTyrrThrPheLysAlaIleIleTyrr	176
QY	292	TATAAAATATCATTCAGACAGCACATGAGCGGGGACACATATAGACAGATCACTAAT	351
Db	177	TyrrLysAsnValIleValThrThrTrpAlaGlySerThrTyrrAlaAlaIleThrAsn	196
QY	352	CGATATACAGATAGACGCCGCGTTTCCTTAACAAGAGATCACGGATCTAATCAGCGCAAA	411
Db	197	GlnTyrrThrAspArgValProValcglyMetGlycyluileThrAspLeuValaspLysLys	216
QY	412	GGAGATGCTCATCTAAAGCAAGATACCTTAGAACAATGTATATCTTGAGCGGTTTGAC	471
Db	217	TrpArgCysLeuSerLysAlaGluTyrrLeuArgSerGlyArgLysValValAlaPheAsp	236
QY	472	AGGGATCGGGAGAAAACAGTAGTACTTCTAAAACCATCAAAATTCACACGCCGCAATCT	531
Db	237	ArgAspAspAspProtpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal	256
QY	532	AGGGCATGGCACAGCACTAATGACACGTATACCGTGTGGGGATCCACATGGATATATCGA	591
Db	257	ArgGlyTrpHisThrThrAspAspValTyrrAlaLeuGlySerAlaGlyLeutyryrArg	276
QY	592	ACGGGAACCTCCGTCAAATTGTATAGAGAAATGGATCGCCGGCTCTGTGTTCCGTAT	651
Db	277	ThrGlyThrSerValAsnCysIleValIGluGluValIGluAlaargSerValtyrProtyr	296
QY	652	TCATATTTTGCATGGCC	669
Db	297	AspSerPheAlaLeuSer	302

RESULT 6

B48474
glycoprotein B - feline herpesvirus 1 (fragment)
C:Species: feline herpesvirus 1

2

C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B48474
R;Spatz, S.J.; Maes, R.K.
Virology 197, 125-136, 1993
A;Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B an
A;Reference number: A48474; MUID:94025539; PMID:8212548
A;Contents: C-27
A;Accession: B48474
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-943 <SPA>
A;Cross-references: GB:S56371; NID:9435792; PIDN:AAB38559.1; PID:9435794
A;Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBIP:138807)
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein

Alignment Scores:	
Pred. No.:	5,486-50
Score:	583.00
Percent Similarity:	56.07%
Best Local Similarity:	42.86%
Query Match:	48.58%
DB:	2
Length:	943
Matches:	120
Conservative:	37
Mismatches:	63
Indels:	60
Gaps:	5

US-09-147-052-3_COPY_1_672 (1-672) x B48474 (1-943)

QY	7	TATTTTAGCGGAATTGCTATATT-----	- 30
Db	22	:::	
QY	31	----	
Db	42	HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61	----
QY	43	ATTCTATAT-----	----
Db	62	ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81	----
QY	70	AGTACCCAAATGTCACATCAAGAAAGTT-----	99
Db	82	ProArgArgThrValAlaIleThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101	----
QY	100	-----GTTTCGAGCGTCACAGTGTCTGAGGAAGAG-----	129
Db	102	ProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121	----
QY	130	TCACGTTTATCTTTGTGCCCCCACCAGTCGGTTCAACCGTGATCCGCTCCTAGAACCGCCG 189	----
Db	122	SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141	----
QY	190	CGAAATGTCGCCACCTTAPGAAAGCCACCAGTGGGTGAAGGAATCGCGATATTATTT 249	----
Db	142	ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161	----
QY	250	AAAGAGATATCAGTCCATATAAATTTAAAGTCAGGCTTTATTATAAAATATCATTCAG 309	----
Db	162	LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181	----
QY	310	ACGACGACATCGACGGGACACATATAGACAGATCAATACATATACAGATAGACG 369	----
Db	182	ThrThrValTrpSerGlySerSerTyrAlaValThrAsnArgTyrThrAspArgVal 201	----
QY	370	CCGTTTCCATTGAAGAGATCAGCGATCTAATCGACGGCAAGAGAGATGCTCATCTAA 429	----
Db	202	ProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeuSerLys 221	----
QY	430	GCAAGATACCTTAGAAACAATGTATATTGTCAGCGTTTCACAGGAGTCCGGGAGAAAA 489	----
Db	222	AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241	----
QY	490	CAAGTACTTCTAAACCATCAAAATTCACACGCCGCCCAATCTAGGCGATGGCACAGACT 549	----
Db	242	GluLeuProLeuLysProProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260	----

RESULT 6

B48474
glycoprotein B - feline herpesvirus 1 (fragment)
C:Species: feline herpesvirus 1

2


```

QY 550 AATGACAGCTATACCGTGTGGGGATCACCATGATATATCCGAACGGAACTCCGTCAT 609
|||||
Db 261 AsnGluThrTyrThrLysLeuValLeuLeuAspPheHisSerGlyThrSerValAsn 280
||| :|||
QY 610 TGTATAGTAGAGAAATGATGCGCGCTCTGTTTCCGTTATTCATATTTGGCAATGGCC 669
|||||
Db 281 CysIleValGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300
|||||

RESULT 7
VGBEBC
glycoprotein g1 precursor - bovine herpesvirus 1 (strain Colorado-1[Cooper-1])
N;Alternate names: glycoprotein 11a; glycoprotein 16; glycoprotein g130; glycoprotein g1
C;Species: bovine herpesvirus 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A28877; A43380
R;Whitbeck, J.C.; Bello, L.J.; Lawrence, W.C.
J. Virol. 62, 3319-3327, 1988
A;Title: Comparison of the bovine herpesvirus 1 g1 gene and the herpes simplex virus typh
A;Reference number: A28877; PMID:88300884; PMID:2841484
A;Accession: A28877
A;Molecule type: DNA
A;Residues: 1-932 <WHI>
A;Cross-references: GB:M21474; NID:g330756; PIDN:AAA46055.1; PID:g330757
R;van Drunen Littell-van den Hurk, S.; Parker, M.D.; Fitzpatrick, D.R.; van den Hurk, J.V
Virology 190, 378-392, 1992
A;Title: Structural, functional, and immunological characterization of bovine herpesviru
A;Reference number: A43380; PMID:92410615; PMID:1326809
A;Accession: A43380
A;Molecule type: protein
A;Residues: 68-87;505-516 <VAN>
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-67/Domain: signal sequence #status predicted <SIG>
F;68-932/Product: glycoprotein g1 #status predicted <GGI>
F;68-504/Product: glycoprotein g1b #status predicted <GIB>
F;505-932/Product: glycoprotein g1c #status predicted <GIC>
F;759-827/Domain: transmembrane #status predicted <TMN>
F;105,153,441,483,640,706/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 3 91e-48 Length: 932
Score: 564.50 Matches: 106
Percent Similarity: 67.96% Conservative: 34
Best Local Similarity: 51.46% Mismatches: 65
Query Match: 47.04% Indels: 1
DB: 1 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x VGBEBC (1-932)

QY 52 GTTACGAATCTATCTCCGATACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTC 111
||| :|||
Db 97 GlyAspAlaAlaSerProAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116
||| :|||
QY 112 CAGTTCCTGAGGAAGCTACGTTTATCTTTGTCGCCACCACGATGGTTCAACCGTG 171
||| :|||
Db 117 GlnAlaAlaGlyGluAsnSerArgPheValCysProProSerGlyAlaThrVal 136
||| :|||
QY 172 ATCCGCTAGAACCGCGCGGAAATGTCCGGAACCTAGAAAGCCAGCGAGTGGGTGAA 231
||| :|||
Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
||| :|||
QY 232 GGAATCCGATATATTAAAGAGAATATCAGTCCCATATAAATTTAAAGTGAGCGCTTAT 291
||| :|||
Db 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIle--- 175
||| :|||
QY 292 TATAAAATATATTCAGACGACGATGACGGGACGACATATAGACAGATCACTAAT 351
||| :|||
Db 176 TyrLysAsnValIleValThrThrTrpAlaGlySerThrTyrAlaIleThrAsn 195
||| :|||
QY 352 CGATATACAGATAGGACCGCGCTTCCATTGAAGACATCAGCATTAATCGACGGCAA 411
||| :|||

```

```

Db 196 GlnTyrThrAspArgValProValGlyMetGlyGluLeuThrAspLeuValAspLysLys 215
QY 412 GGAAGATGCTCATCTAAAGCAAGATACCTTTAGAAACAATATATATGTTGAAGCGTTTGAC 471
|||||
Db 216 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgValValAlaPheAsp 235
|||||
QY 472 AGCGATCGGGGAGAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCT 531
|||||
Db 236 ArgAspAspProTrpGluAlaProLeuLysPheAlaArgLeuSerAlaProGlyVal 255
|||||
QY 532 AGGCGATGGGCACACGACTATAGAGCTATACCGTGTGGGATCACCATCGATATATCGA 591
|||||
Db 256 ArgGlyTrpHisThrAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 275
|||||
QY 592 ACGGGAACCTCCGCTCAATTGTATAGTAGAGGAATGGATCGCGCTGTGTGTTCCGTAT 651
|||||
Db 276 ThrGlyThrSerValAsnCysIleValGluValGluAlaArgSerValTyrProTyr 295
|||||
QY 652 TCATATTTTCCAATGGCC 669
|||||
Db 296 AspSerPheAlaLeuSer 301
|||||

RESULT 8
JH0109
glycoprotein 14 precursor - equine herpesvirus 1
C;Species: equine herpesvirus 1
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 26-Aug-1999
C;Accession: JH0109
R;Guo, P.
Gene 87, 249-255, 1990
A;Title: Characterization of the gene and an antigenic determinant of equine herpesvi
A;Reference number: JH0109; PMID:90236317; PMID:1692002
A;Accession: JH0109
A;Molecule type: DNA
A;Residues: 1-979 <GUO>
A;Cross-references: GB:M34861; NID:g330900; PIDN:AAA46086.1; PID:g330901
A;Note: glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies,
C;Genetics:
A;Gene: gpl4
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-86/Domain: signal sequence #status predicted <SIG>
F;87-979/Product: glycoprotein 14 #status predicted <MAT>
F;831-871/Region: hydrophobic
F;165,275,380,423,497,514,560,727,749,951,970/Binding site: carbohydrate (Asn) (coval
Alignment Scores:
Pred. No.: 4 97e-48 Length: 979
Score: 563.50 Matches: 113
Percent Similarity: 59.36% Conservative: 36
Best Local Similarity: 45.02% Mismatches: 69
Query Match: 46.96% Indels: 33
DB: 2 Gaps: 4

US-09-147-052-3_COPY_1_672 (1-672) x JH0109 (1-979)

QY 13 AGCGGAATTCATATTTTCTTATATGTTATCTATATGTT----- 54
||| :|||
Db 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
||| :|||
QY 55 -----ACGAACATCTCCGAGTACCCAAATGTGACATCA----- 90
||| :|||
Db 86 ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis 105
||| :|||
QY 91 -----AGAGAAGTGTGT 102
||| :|||
Db 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 135
||| :|||
QY 103 TCGAGCGTCAGTGTCTGAGGAAGAGTACCTTTTATCTTTTATCTTTTCCCGCCAGCGGGT 162
||| :|||
Db 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145
||| :|||
QY 163 TCAACCGTGATCGTCTAGAACCGCGCGGAAAATGTCGGAACCTAGAAAGCCACCGAG 222
||| :|||

```

```

Db      146 SerThrValAlaArgIleGluProArgThrCysProLysPheAspLeuGlyArgAsn 165
      223 TGGGGTGAAGGATCGCGATATTATTTAAAGAGAAATATCAGTCCCATATAAATTTAAAGTG 282
      166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla 185
      283 ACGCTTTATTATAAAATATCATTCAGACGACGACATCGCGGGGACACATATATAGACAG 342
      186 AsnValTyrTyrLysAspIleValThrArgValTyrLysGlyTyrSerHisThrSer 205
      343 ATCACTAATCGATATACAGATAGGACGCCGTTTCCATTTGAAGAGATCACGGATCTAATC 402
      206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuIle 225
      403 GAGCGCAAGGAGATGCTCATCTAAACACAGATACCTTAAACCAATATGTTGTTGAA 462
      226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMethHis 245
      463 CGCTTTGACAGGATCGCGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACAG 522
      246 AlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
      523 CCCGAATCTAGCGATGCGACACGACTAATGAGAGC-----TATACCGTGTGGGATCA 576
      266 ProGlyAlaArgAlaTyrThrSerThrSerValAsnCysIleValGluGluValAlaArg 304
      577 CCATGGATATATCAAGCGGACCTCCGTCATTTGTTGCAATGCC 669
      285 SerValTyrProTyrAspSerPheAlaLeuSer 315

```

```

RESULT 9
VGBEC6
glycoprotein B precursor - equine herpesvirus 1 (strain Ab4p)
A:Note: host Equus caballus (domestic horse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
R:Accession: G36798
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Submitted to GenBank, March 1992
A:Description: The DNA sequence of equine herpesvirus-1.
A:Reference numbers: A36805
A:Accession: G36798
A:Molecule type: DNA
A:Residues: 1-980 <TEU>
A:Cross-references: GB:M86664; NID:g330791; PIDN:AAH02468.1; PID:g330825
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 33
Pred. No.: 4,97e-48 Length: 980
Score: 563.50 Matches: 113
Percent Similarity: 59.36% Conservative: 36
Best Local Similarity: 45.02% Mismatches: 69
Query Match: 46.96% Indels: 33
Gaps: 1
Alignment Scores:

```

```

US-09-147-052-3_COPY_1_672 (1-672) x VGBEC6 (1-980)
QY      13 AGGGGAATTCATATTTTCTTATAGTATTATCTATAGT----- 54
Db      66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
QY      55 -----ACGAACTCATCTCCGAGTACCCAAATATGACATCA----- 90
Db      86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105
QY      91 -----AGACAAGTTTCTT----- 102
Db      106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY      103 TCGAGCGTCCAGTTGCTCTGAGGAAGAGTCTACGTTTTTATCTTTTCCCGCCACGAGTGGT 162
Db      126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145
QY      163 TCACCGGTGATCGCTCTAGAACCGCGCGGAAATGTCCGAACTCAGTAAAGCCACCGAG 222
Db      146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY      223 TGGGGTGAAGCAATCGCATATTATTTAAAGAGATATCAGTCCCATATAAATTTAAAGTG 282
Db      166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla 185
QY      283 ACGCTTTATTATAAAATATCATTCAGACGACGACATCGCGGGGACACATATAGACAG 342
Db      186 AsnValTyrTyrLysAspIleValThrArgValTyrLysGlyTyrSerHisThrSer 205
QY      343 ATCACTAATCGATATACAGATAGGACGCCGTTTCCATTTGAAGAGATCACGGATCTAATC 402
Db      206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuIle 225
QY      403 GAGCGCAAGGAGATGCTCATCTAAACACAGATACCTTAAACCAATATGTTGTTGAA 462
Db      226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMethHis 245
QY      463 CGCTTTGACAGGATGCGGGAGAAAACAAGTACTTCTAAACCATCAAAATTTCAACAG 522
Db      246 AlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
QY      523 CCCGAATCTAGCGATGCGACACGACTAATGAGAGC-----TATACCGTGTGGGATCA 576
Db      266 ProGlyAlaArgAlaTyrThrSerThrSerValAsnCysIleValGluGluValAlaArg 304
QY      577 CCATGGATATATCAAGCGGACCTCCGTCATTTGTTGCAATGCC 669
Db      285 ProThrArgHisTyrThrSerThrSerValAsnCysIleValGluGluValAlaArg 304
QY      637 TCTGTGTTTCCGTTATTCATATTTTTCATATGTTGCAATGCC 669
Db      305 SerValTyrProTyrAspSerPheAlaLeuSer 315

```

```

RESULT 10
VGBEQH
glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C:Species: equine herpesvirus 4
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A31880
R:Riggio, M.P.; Cullinane, A.A.; Onions, D.E.
J. Virol. 63, 1123-1133, 1989
A:Title: Identification and nucleotide sequence of the glycoprotein gB gene of equine
A:Reference number: A31880; MUID:89125704; PMID:2915378
A:Accession: A31880
A:Molecule type: DNA
A:Residues: 1-919 <RCG>
A:Cross-references: GB:M26171; NID:g341446; PIDN:AAA46106.1; PID:g514920
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-919/Product: glycoprotein B #status predicted <GPB>

```

F:740-809/Domain: transmembrane #status predicted <TM>

F:106,216,321,364,438,456,493,499,666,688/Binding site: carbohydrate (Asn) (covalent) #s

Alignment Scores:

Pred. No.:	1,24e-47	Length:	919
Score:	559,50	Matches:	109
Percent Similarity:	64.73%	Conservative:	36
Best Local Similarity:	48.66%	Mismatches:	58
Query Match:	46.62%	Indels:	21
DB:	1	Gaps:	4

US-09-147-052-3_COPY_1_672 (1-672) x VGBE0H (1-919)

QY	58	AACATCTCTCCGAGTACC-----	CAAAATGTG 84
DB	34	SerSerGlnProSerThrProAlaSerThrGlnSerAlaLysThrValAspGlnThrLeu 53	:
QY	85	ACATCAAGAGAAGTGTTCGAGCGTCCAGTCTCTGAGGAAGAGTCT-----	132
DB	54	LeuProThrGlnThrProAspProLeuArgLeuAlaValArgLysGlyIleLeuAla 73	:
QY	133	-----ACGTTTATCTTCTCCCGCCACCGAGTGGTTCAACCGTGATCCGCTCAGAA 183	:
DB	74	GluAspGlyAspPheThrCysProProThrGlySerThrValValArgIleGlu 93	:
QY	184	CCGCGCGGAAATGTCCCGAACCTAGNAAGCCACCGAGTGGGTGAAGAAATCGGATA 243	:
DB	94	ProProArgSerCysProLysPheAspLeuGlyArgAsnPheThrGluGlyIleAlaVal 113	:
QY	244	TTATTTAAAGAGAAATATCATCTCCATATAATTTAAAGTGACGCTTTATATAAAATATC 303	:
DB	114	IlePheLysGluAsnIleAlaProTyrLysPheArgAlaAsnValTyrTyrLysAspIle 133	:
QY	304	ATTCAGACGACATATGACGGGACGACATATAGACAGATCACTAATCGATATACAGAT 363	:
DB	134	ValValThrLysValTyrLysGlyTyrSerHisThrSerLeuSerAspArgTyrAsnAsp 153	:
QY	364	AGGACGCGCTTTCATTTGAAGAGATCACCGATCTAATCCACGCGCAAGAAAGATGCTCA 423	:
DB	154	ArgValProValSerValGluGluIlePheThrLeuIleAspSerLysGlyLysCysSer 173	:
QY	424	TCATAAGCAAGATACCTTAGAAACAATGTATATGTTGACGCGTTTGACAGGGATCGGGA 483	:
DB	174	SerLysAlaGluTyrLeuArgAspAsnIleMetHisAlaTyrHisAspAspGluAsp 193	:
QY	484	GAATAACAAGTACTTCTAAACCATCAAAATTCACACCGCCGGAATCTAGGCGATGGCAC 543	:
DB	194	GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaArgAlaTyrGln 213	:
QY	544	ACGACTAATGAGAGC-----TATACCGTGTGGGATCACCATGGATATATCAACGGGA 597	:
DB	214	ThrThrAsnAspThrThrSerTyrValGlyTyrPheMetProTyrArgHisTyrThrSer 232	:
QY	598	ACCTCCGCTCAATGTATAGTAGAGAAATGGATCCCGCTCTGTGTTCCGTTATTCATAT 657	:
DB	233	ThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspSer 252	:
QY	658	TTTGCAATGCC 669	:
DB	253	PheAlaLeuSer 256	:

RESULT 11

T42576

probable envelope protein 33 - equine herpesvirus 4 (strain NS80567)

C:Species: equine herpesvirus 4

A:Variety: strain NS80567

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T42576

R:Gelford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.

J. Gen. Virol. 79, 1197-1203, 1998

A:Title: The DNA sequence of equine herpesvirus-4.

A:Reference number: 222173; MUID:98264497; PMID:9603335

A:Accession: T42576

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-975 <TEL>

A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59549.1; PID:g2605977

A:Experimental source: strain NS80567

C:Genetics:

A:Gene: 33

C:Superfamily: herpesvirus glycoprotein B

Alignment Scores:

Pred. No.:	1,25e-47	Length:	975
Score:	559,50	Matches:	109
Percent Similarity:	64.73%	Conservative:	36
Best Local Similarity:	48.66%	Mismatches:	58
Query Match:	46.62%	Indels:	21
DB:	2	Gaps:	4

US-09-147-052-3_COPY_1_672 (1-672) x T42576 (1-975)

QY	58	AACATCTCTCCGAGTACC-----	CAAAATGTG 84
DB	90	SerSerGlnProSerThrProAlaSerThrGlnSerAlaLysThrValAspGlnThrLeu 109	:
QY	85	ACATCAAGAGAAGTGTTCGAGCGTCCAGTCTCTGAGGAAGAGTCT-----	132
DB	110	LeuProThrGlnThrProAspProLeuArgLeuAlaValArgLysGlyIleLeuAla 129	:
QY	133	-----ACGTTTATCTTCTCCCGCCACCGAGTGGTTCAACCGTGATCCGCTCAGAA 183	:
DB	130	GluAspGlyAspPheThrCysProProThrGlySerThrValValArgIleGlu 149	:
QY	184	CCGCGCGGAAATGTCCCGAACCTAGNAAGCCACCGAGTGGGTGAAGAAATCGGATA 243	:
DB	150	ProProArgSerCysProLysPheAspLeuGlyArgAsnPheThrGluGlyIleAlaVal 169	:
QY	244	TTATTTAAAGAGAAATATCATCTCCATATAATTTAAAGTGACGCTTTATATAAAATATC 303	:
DB	170	IlePheLysGluAsnIleAlaProTyrLysPheArgAlaAsnValTyrTyrLysAspIle 189	:
QY	304	ATTCAGACGACATATGACGGGACGACATATAGACAGATCACTAATCGATATACAGAT 363	:
DB	190	ValValThrLysValTyrLysGlyTyrSerHisThrSerLeuSerAspArgTyrAsnAsp 209	:
QY	364	AGGACGCGCTTTCATTTGAAGAGATCACCGATCTAATCCACGCGCAAGAAAGATGCTCA 423	:
DB	210	ArgValProValSerValGluGluIlePheThrLeuIleAspSerLysGlyLysCysSer 229	:
QY	424	TCATAAGCAAGATACCTTAGAAACAATGTATATGTTGACGCGTTTGACAGGGATCGGGA 483	:
DB	230	SerLysAlaGluTyrLeuArgAspAsnIleMetHisAlaTyrHisAspAspGluAsp 249	:
QY	484	GAATAACAAGTACTTCTAAACCATCAAAATTCACACCGCCGGAATCTAGGCGATGGCAC 543	:
DB	250	GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaArgAlaTyrGln 269	:
QY	544	ACGACTAATGAGAGC-----TATACCGTGTGGGATCACCATGGATATATCAACGGGA 597	:
DB	270	ThrThrAsnAspThrThrSerTyrValGlyTyrPheMetProTyrArgHisTyrThrSer 288	:
QY	598	ACCTCCGCTCAATGTATAGTAGAGAAATGGATCCCGCTCTGTGTTCCGTTATTCATAT 657	:
DB	289	ThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspSer 308	:
QY	658	TTTGCAATGCC 669	:
DB	309	PheAlaLeuSer 312	:

RESULT 12

B48349

glycoprotein B precursor - ateline herpesvirus 1 (strain Lennette)

C:Species: ateline herpesvirus 1

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994

C:Accession: B48349

R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A:Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine
A:Reference number: A48349; MUID:93228440; PMID:8385913
A:Accession: B48349
A:Molecule type: DNA
A:Residues: 1-933 <EBE>
A:Note: sequence extracted from NCBI backbone (NCBIN:129063, NCBI:P:129065)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-933/Product: glycoprotein B #status predicted <MAT>
F:775-794/Domain: transmembrane #status predicted <TM1>
F:801-818/Domain: transmembrane #status predicted <TM2>
F:107,161,418,450,697,747/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1,26e-44 Length: 933
Score: 529.50 Matches: 101
Percent Similarity: 62.62% Conservative: 28
Best Local Similarity: 49.03% Mismatches: 74
Query Match: 44.12% Indels: 3
DB: 1 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x B48349 (1-933)

QY	52	GGTACGACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTGTTTCGAGCGTC	111
DB	108	GlyThrArgSerAlaAlaArgGln-----LeuArgGluSerLeuArgArgile	124
QY	112	CAGTTGTCTGAGGAAGAGTACGTTTATCTTGTCTCCGCCACCAAGTGGTTCACCGTG	171
DB	125	GlnAlaGluTyrAlaAlaSerAlaPheTyrValCysProProThrGlyAlaThrVal	144
QY	172	ATCCGTCTAGACCCCGCGAAATGTCGCCAACCCTAGAAAGCCACCGAGTGGGTCAA	231
DB	145	ValGlnPheGluGluProArgProCysProAspValAlaAlaGlyLysAsnPhetrGlu	164
QY	232	GGAATCGCGATATATTTAAAGAGATATCATGCTCATATAATTTAAAGTCAGCGTTAT	291
DB	165	GlylleAlaValAlaPheLysGluAsnIleAlaProTyrLysPheThrAlaThrLysTyr	184
QY	292	TATAAATATCATTCAGACGACATGCGGCGGACGACATATAGACATCACTAAT	351
DB	185	TyrLysGluThrValSerGlnThrTrpGlnGlySerArgTyrLeuGlnLeuThrGly	204
QY	352	CGATACAGATAGGAGCGCCCTTCCATTGAAGAGATCACGGATCTAATCGACGCCAAA	411
DB	205	LeuTyrAsnAspArgAlaProValProPheSerGluIleThrAspLeuIleAsnGlyLys	224
QY	412	GGAAGTGTCTATCTAAAGCAAGATACCTTAGAACATGTATGTGACGGTTGAC	471
DB	225	GlyArgCysArgSerAspValThrTyrThrArgSerGlnArgValThrAlaTyrAsp	244
QY	472	AGGATGCGGCAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCGCAATCT	531
DB	245	GlyAspGluTrpGlyArgGluValAlaLeuValProAlaLysThrSerThrProAsnSer	264
QY	532	AGGCGATGCGACACGACTAATAGACGTATACCGTGTGGGGATCCACCATGATATCGA	591
DB	265	ArgGlyTrpTyrThrThrAspArgValTyrAlaProAsnAlaHisAlaGlyPheTyrLys	284
QY	592	ACGGGAACCTCCCTCAATGTATAGTAGGAATGATGCCGCTGTGTTCCGGTAT	651
DB	285	ThrGlyThrThrValAsnCysIleValGluGluMetGluAlaArgSerAlaPheProTyr	304
QY	652	TCATATTTTGCATGGCC 669	
DB	305	AspSerPheValLeuAla 310	

RESULT 13
VGBE2H
glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)

C:Species: equine herpesvirus 1
C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
C:Accession: A31241; J00008
R:Whalley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Bell, C.W.; Woodworth, L
J. Gen. Virol. 70, 383-394, 1989
A:Title: Identification and nucleotide sequence of a gene in equine herpesvirus 1 ana
A:Reference number: A31241; MUID:89279217; PMID:2543744
A:Accession: A31241
A:Molecule type: DNA
A:Residues: 1-980 <WHA>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-85/Domain: signal sequence #status predicted <SIG>
F:70-78/Domain: transmembrane #status predicted <TM1>
F:86-980/Product: glycoprotein B #status predicted <MAT>
F:853-968/Domain: transmembrane #status predicted <TM2>
F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (c

Alignment Scores:
Pred. No.: 3.2e-44 Length: 980
Score: 525.50 Matches: 109
Percent Similarity: 57.94% Conservative: 37
Best Local Similarity: 43.25% Mismatches: 71
Query Match: 43.79% Indels: 35
DB: 1 Gaps: 6

US-09-147-052-3_COPY_1_672 (1-672) x VGBE2H (1-980)

QY	13	AGCGGAATTCATATATTTTCTTATATGTTATCTATATGTT	54
DB	66	ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla	85
QY	55	-----ACGAACCTCATCTCGAGTACCCAAATGTGACATCA	90
DB	86	ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis	105
QY	91	-----AGAGAAAGTTGTT	102
DB	106	GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal	125
QY	103	TCGACGCTCCAGTTGTCGTAGGAAGAGTCTACGTTTATCTTCTCCGCCACCGAGTGGT	162
DB	126	ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly	145
QY	163	TCACCGTGATCCCTCTAGAACCCCGCGAAATGTCGCCAACCCTAGAAAGCCACCGAG	222
DB	146	SerThrValValArgGileGluProProArgThrCysProLysPheAspLeuGlyArgAsn	165
QY	223	TGGGTGAAGGAATCGCGATATTTAAAGAGAAATATCATGCTCATATAATTTAAAGTG	282
DB	166	PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla	185
QY	283	ACGCTTTATTTATAAATATCATTCAGACGACATGCGGCGGACGACATATAGACAG	342
DB	186	AsnValTyrTyrLysAspIleValThrArgValTyrLysGlyTyrSerHisThrSer	205
QY	343	ATCCTAATTCATATACAGATAGGACGCCGTTTCCATTTGAAGATCACGGATCAATC	402
DB	206	LeuSerAspArgTyrAsnAspArgValProValSerValGluIlePheGlyLeuIle	225
QY	403	GACGCCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACATGTATATGTTGAA	462
DB	226	AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMethHisHis	245
QY	463	GGCTTTCAGCGGACCGGAGGAAAAACAAGTA---CTTCTAAACCATCAAAATTCAC	519
DB	246	AlaTyrHisAspAspGluAspGluValGluLeuAspLeuCysArgProSer---LeuGln	264
QY	520	ACGCCCGATCTAGGCGATGGCACACGACTAATAGACG-----TATACCGTGTGGGA	573
DB	265	LeuArgGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyrValGlyTrp---	283
QY	574	TCACCATGATATATCGAACGGAACCTCCGTCATTTGTTATGATAGAGAAATGATGCC	633

```
Db 284 MetProTyrArgHisTyrThrSerThrSerValAsnCysIleValGluGluAla 303
||||| ... ||| |||||||||||||||||||||||||||||||||||
QY 634 CGCTCTGTGTTCCGTCATTTCATATTGCAATGGCC 669
|||||...:|||||
Db 304 ArgSerValTyrProTyrAspSerPheAlaLeuSer 315
|||||...:|||||
RESULT 14
D48349
glycoprotein B precursor - simian herpesvirus 1 (strain MV-5-4-PSL)
C:Species: simian herpesvirus 1
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
C:Accession: D48349
R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A:Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine m
A:Reference number: A48349; MUID:93228440; PMID:8385913
A:Accession: D48349
A:Molecule type: DNA
A:Residues: 1-920 <BE>
A>Note: sequence extracted from NCBI backbone (NCBIN:129066, NCBIP:129068)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-920/Product: glycoprotein B #status predicted <MAT>
F:576-592/Region: hydrophobic
F:761-780/Domain: transmembrane #status predicted <TM1>
F:787-804/Domain: transmembrane #status predicted <TM2>
F:98,119,152,409,441,683,733/Binding site: carbohydrate (Asn) (covalent) #status predict
Alignment Scores:
Pred. No.: 4,48e-44 Length: 920
Score: 524.00 Matches: 96
Percent Similarity: 62.75% Conservative: 32
Best Local Similarity: 47.06% Mismatches: 76
Query Match: 43.67% Indels: 0
DB: 1 Gaps: 0
US-09-147-052-3_COPY_1_672 (1-672) x D48349 (1-920)
QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAAGTTGTTTCGAGCGTCAGTGG 117
||| ... ||| |||...:|||||
Db 98 AsnLysThrAlaGluAspAlaArgAlaGlnLeuArgGlnSerValArgGlnIleArgAla 117
||| ... ||| |||...:|||||
QY 118 TCTGAGAGAGCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTGATCCGT 177
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 118 GluAsnAlaThrSerMetPheTyrValCysProProThrGlyAlaThrValValGln 137
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 178 CTAGAACCCGCGGAAATGTCCTCCGAACTAGAAAGCCACCGAGTGGGGTGAAGGAATC 237
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 138 PheGluLupProArgProCysProAspValAlaAlaGlyLysAsnPheThrGluGlyIle 157
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 238 GCGATATATTTAAAGAGATATTCAGTCCATATAAATTTAAAGTACGCTTTATTATAA 297
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 158 AlaValIlePheLysGluAsnIleAlaProTyrLysPheThrAlaThrMetTyrTyrLys 177
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 298 AATATCATTCAGACGACATCGGCGGACATATAGACATCACTAATCATAT 357
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 178 GluIleThrValThrGlnThrTrpGlnGlySerArgTyrLeuGlnLeuThrGlyLeuTyr 197
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 358 ACAGATAGGACCGCTTCCATTGAAGATCACCAGTCAATCCAGCGCAAGGAAGA 417
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 198 AsnAspArgAlaProValProPheGluGluIleThrAspValIleAsnAlaLysGlyLeu 217
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 418 TGTCTATCTAAAGCAAGATACCTTAGAACAAATGTATGTGAACGCTTTGACAGGAT 477
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 218 CysArgSerAspValThrTyrValArgSerGlnArgValThrAlaTyrAspArg 237
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 478 GCGGGAGAAAACAGTACTTCTPAAACCATCAAAATTCACACGCCCGCAATAGGGCA 537
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 238 GluTrpGlyArgGluValLysLeuValProSerLysThrSerThrProAsnSerArgGly 257
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 538 TGGCACACGACTAATCAGACGCTATACCGTGTGGGGATCACCATCGGATATATCGAACGGGA 597
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
```

```
Db 258 TrpTyrThrThrAspArgMetTyrAlaProAsnAlaHisAlaGlyPheTyrLysAlaGly 277
|||||...:|||||...:|||||...:|||||...:|||||...:|||||
QY 598 ACCTCCGTCATATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGTCATTTCATAT 657
|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 278 ThrThrValAsnCysIleValGluGluValGluAlaArgSerAlaTyrProTyrSerAsn 297
|||||...:|||||...:|||||...:|||||...:|||||...:|||||
QY 658 TTTGCAATGGCC 669
||| ...:|||||
Db 298 PheValLeuAla 301
||| ...:|||||
RESULT 15
VBESA
glycoprotein B precursor - simian herpesvirus SA8 (strain B264)
C:Species: simian herpesvirus SA8
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: JQ1332
R:Borchers, K.; Weigelt, W.; Buhk, H.J.; Ludwig, H.; Mankertz, J.
J. Gen. Virol. 72, 2299-2304, 1991
A:Title: Conserved domains of glycoprotein B (gB) of the monkey virus, simian agent 8
A:Reference number: JQ1332; MUID:91374035; PMID:1895066
A:Accession: JQ1332
A:Molecule type: DNA
A:Residues: 1-885 <BE>
A:Cross-references: EMBL:X56935; NID:g60438; PIDN:CAA40256.1; PID:g60439
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-885/Product: glycoprotein B #status predicted <MAT>
F:712-732/Domain: transmembrane #status predicted <TM1>
F:737-752/Domain: transmembrane #status predicted <TM2>
F:760-780/Domain: transmembrane #status predicted <TM3>
F:68,122,379,411,659/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:97-558,114-514,188-252,345-393,581-618/Disulfide bonds: #status predicted
Alignment Scores:
Pred. No.: 4,49e-42 Length: 885
Score: 504.00 Matches: 94
Percent Similarity: 62.75% Conservative: 34
Best Local Similarity: 46.08% Mismatches: 76
Query Match: 42.00% Indels: 0
DB: 1 Gaps: 0
US-09-147-052-3_COPY_1_672 (1-672) x VBESA (1-885)
QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAAGTTGTTTCGAGCGTCAGTGG 117
||| ... ||| |||...:|||||
Db 68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleLysAla 87
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 118 TCTGAGAGAGCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTGATCCGT 177
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 88 ArgAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValValGln 107
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 178 CTAGAACCCGCGGAAATGTCCTCCGAACTAGAAAGCCACCGAGTGGGGTGAAGGAATC 237
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 108 PheGluLupProArgProCysProAspValAlaAlaGlyLysAsnPheThrGluGlyIle 127
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 238 GCGATATATTTAAAGAGATATTCAGTCCATATAAATTTAAAGTACGCTTTATTATAA 297
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 128 AlaValIlePheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 147
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 298 AATATCATTCAGACGACATCGGCGGACATATAGACATCACTAATCATAT 357
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 167
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 358 ACAGATAGGACCGCTTCCATTGAAGATCACCAGTCAATCCAGCGCAAGGAAGA 417
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 168 GluAspArgAlaProValProPheGluGluValIleAsnAlaLysGlyVal 187
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
QY 418 TGTCTATCTAAAGCAAGATACCTTAGAACAAATGTATGTGAACGCTTTGACAGGAT 477
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
Db 188 CysArgSerThrAlaLysTyrValArgAsnMetGluSerThrAlaPheHisArgAsp 207
||| ...:|||||...:|||||...:|||||...:|||||...:|||||
```

Qy	478	GCGGGAGAAACAGTACTTCTAAACCATCAAAATTCACACGCGCCGAATCTAGGGCA	537
Db	208	AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly	227
Qy	538	TGGCACACGACTAATGAGACGTATACCGTGTGGGGATCACCATGGATATATCGAACGGGA	597
Db	228	TriHisThrThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly	247
Qy	598	ACCTCCGTCATTTGTATAGTAGAGAAATGGATGCCGCTCTGTGTTTCGGTATTCATAT	657
Db	248	ThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspGlu	267
Qy	658	TTTGCAATGGCC	669
Db	268	PheValLeuAla	271

Search completed: October 8, 2003, 18:18:46
 Job time : 31.1621 secs

118	TCTGAGGAAGAGTCTACGTTTATCTTTGTGTCCTCCACCACGAGTGGGTTCAACCGTATCCGT	177
Qy		
119		
Db		
120	GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValGln	121
Qy		
121		
Db		
122	CTGAACCGCGCGGAAATGTCGCCAACCTAGAAAGCCACCGAGTGGGGTGAAGGAATC	237
Qy		
123		
Db		
124	PheGluGlnProArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyLe	141
Qy		
125		
Db		
126	CGCATATTTAAAGACAAATACAGTCCATATAATTTAAAGTACGCGTTTATATATAA	297
Qy		
127		
Db		
128	AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys	161
Qy		
129		
Db		
130	AAATATCATTCAGACGACGACATGGAGGGGCGACATATAGACGATCACTAATCGATAT	357
Qy		
131		
Db		
132	AspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGlyLePhe	181
Qy		
133		
Db		
134	ACATATAGGACGCCGTTTCATTTGAACGATCACGGATCTAATCGCGGCAAGGAAGA	417
Qy		
135		
Db		
136	GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnThrLysGlyVal	201
Qy		
137		
Db		
138	TGCCTCATCTAAAGCAAGATACCTTAGAACCAATGTATGTGTGAACGCTTTCACAGGGAT	477
Qy		
139		
Db		
140	CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgAsp	221
Qy		
141		
Db		
142	CGCGGAGAAAAACAAGTACTCTTAAACCATCAAAATTCACACGCCCGCAATCTAGGGCA	537
Qy		
143		
Db		
144	AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArgGly	241
Qy		
145		
Db		
146	TGGCACACGACTAATGACGACTATACCGTGTGGGATCACCATGGATATATCGACGGGA	597
Qy		
147		
Db		
148	TrpHisThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly	261
Qy		
149		
Db		
150	ACCCTCCGTCATTTGATAGAGAAATGGATGCCGCTCTGTCTTTCGCTATTCATAT	657
Qy		
151		
Db		
152	ThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspGlu	281
Qy		
153		
Db		
154	TTTGCAATGGCC	669
Qy		
155		
Db		
156	PheValLeuAla	285
Qy		

```

RESULT 2
US-09-350-841A-1589
; Sequence 1589, Application US/09350841A
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS F
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS,
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350, 841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1589
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Human
US-09-350-841A-1589

```

Alignment Scores:	
Pred. No.:	1,116-12
Score:	201.50
Percent Similarity:	43.61%
Best Local Similarity:	27.75%
Query Match:	16.79%
DB:	5
Length:	854
Matches:	63
Conservative:	36
Mismatches:	109
Indels:	19
Gaps:	6

US-09-147-052-3_COPY_1_672 (1-672) x US-09-350-841A-1589 (1-854)

Qy 16 CGGAATTGCATATTTTCTCTATAGTTATTCTATATGGT----- 54
||| :: : : : : |||
Db 3 ArgArgValLeuSerValValValLeuAlaLeuAlaCysArgLeuGlyAla 22
Qy 55 ----ACGACTCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTC 111

Db	23	GlnThrProGluGlnProAlaProAlaThrThrValGlnProThrAlaThrArgGln	42
QY	112	CAGTTGCTGAGGAGAGACTCTACGTTTATCTTCTGCCCCACCACTGGGTTCAACCGTG	171
Db	43	GlnThrSer-----PheProPheArgValCysGluLeuSerSerHisGlyAspLeu	59
QY	172	ATCCGTCTAGAACCGCCGCGAAAATGTCCCGAACCTAGAAAAGCCACCGAGTGGGTGAA	231
Db	60	PheArgPheSerSerAspIleGlnCysProSerPheGlyThrArgGluAsnHisThrGlu	79
QY	232	GGATCCGCGATATTATTAAAGACAGATATCATGTCATATATAATTTAAAGTGCAGCTTAT	291
Db	80	GlyLeuLeuMetValPheLysAspAsnIleIleProTyrSerPheLysValArgSerTyr	99
QY	292	TATAAAATATCATTCACAGCACACATCGGCGGGGACGACATATAGACAGATCACTAAT	351
Db	100	ThrLysIleValThrAsnIleLeuIleTyrAsnGlyTyrTyrAlaAspSerValThrAsn	119
QY	352	CGATATACAGATAGACGCCCGTTCCATTGAGAGATCAGCGATCTATCGACGGCGAAA	411
Db	120	ArgHisGluGluLysPheSerValAspSerTyrGlu--ThrAspGlnMetAspThrIle	138
QY	412	GGAAGATGCTCATCTAAAGCAAGATACCTTAGAAC-----AATGTATATCTGAA	462
Db	139	TyrGlnCysTyrAsnAlaValLysMetThrLysAspGlyLeuThrArgValTyrVal---	157
QY	463	CGGTTTGACAGGGATCGGGGAGAAAAACAAGTACTTCTAAAACCATCAAAATTCACACG	522
Db	158	-----AspArgAspGlyValAsnIleThrValAsnLeuLysProThrGlyGlyLeuAla	175
QY	523	CCCGAAATCTAGGGCATGCACACGACTAATGAGACGTATACCGTGTGGGGA---TCACCA	579
Db	176	AsnGlyValArgTyrAlaSerGlnThrGluLeuTyrAspAlaProGlyTyrLeuIle	195
QY	580	TGGATATATCGAACCGGAACCTCCGTCAATTGTATAGTAGAGGAATGGATGCCGCCTCT	639
Db	196	TrpThrTyrArgThrArgThrValAsnCysLeuIleThrAspMetMetAlaLysSer	215
QY	640	GTGTTTCGGTATTCATATTTT	660
Db	216	AsnSerProPheAspPhePhe	222

```

RESULT 3
US-10-425-114A-42745
; Sequence 42745, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecu
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,1
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42745
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700471069
US-10-425-114A-42745

```

Alignment Scores:	
Pred. No.:	1.57
Score:	82.00
Percent Similarity:	33.67%
Best Local Similarity:	20.92%
Query Match:	6.83%
Length:	218
Matches:	41
Conservative:	25
Mismatches:	56
Indels:	74

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

[illegible][illegible]

RESULT 5
PCT-US03-38227-5096
; Sequence 5096, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Tolnetta A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PIRTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott E.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSIIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIXOU, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT

```

US-10-425-114A-64009
; Sequence 64009, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: KOVALIC, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64009
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3245-288-C12_FLI pep
US-10-425-114A-64009

Alignment Scores:
Pred. No.: 1.85 Length: 434
Score: 82.00 Matches: 41
Percent Similarity: 33.67% Conservative: 25
Best Local Similarity: 20.92% Mismatches: 56

```

```

; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5096
; LENGTH: 1130
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 930177.PT544p
PCT-US03-28227-5096

```

```

Alignment Scores:
Pred. No.: 2 93 Length: 1130
Score: 81.00 Matches: 42
Percent Similarity: 35.38% Conservative: 33
Best Local Similarity: 19.81% Mismatches: 77
Query Match: 6.75% Indels: 60
DB: 1 Gaps: 10

```

US-09-147-052-3_COPY_1_672 (1-672) x PCT-US03-28227-5096 (1-1130)

```

QY 67 CCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTCCGAGCGTCCAGTTGCTGAGGAA 126
Db 141 ProSerArgSerProLeuThrPheTrpThrLysGluLysLeuGluProIleThrLeuGlnSerGly 160
QY 127 GAGTCTACGTTTATCTTTGTCCTCCACAGGAGTGGTTCACCGTATCGGTGTAGAACGG 186
Db 161 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 175
QY 187 CCG-----CGAAATGTCGCGAACCTAGAAAGCCACC 219
Db 176 ProIleIlePheTrpMetAspAsnSerPheGlnArgLeuProGlnSerGluArgValSer 195
QY 220 GAGTGGGTGAAGGAATCCGATATATTAAAGAGAATATCAGTCCCATATAAATTTAA 279
Db 196 GlnGlyLeuAsnGly----- 200
QY 280 GTGAGCTTATATAAATATCAT---CAGACGACGACATGACGGGGGACGACATAT 336
Db 201 ---AspLeuTyPheSerAsnValLeuProGluAspThrArgGluAspTyrIleCystyr 219
QY 337 AGACAGATCATAATCGATATACA-----GATAGAGCGCGGTTTCCATTGGAAGAGATC 390
Db 220 AlaArgPheAsnHisThrGlnThrIleGlnGlnLysGlnProIleSerValIleValIle 239
QY 391 ACGGATCTAATCGACGCGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTGAACCAAT 450
Db 240 Ser-----ValAsp-----GluLeuAsnAspThr 247
QY 451 GTATATGTTGAAGCGTTTGACAGGAT---CGGGAGAAACAAAGTACTTCTAAACCA 507
Db 248 IleAlaAlaAsnLeuSerAspThrGluPheTyrGlyAlaLysSerArgGluArgPro 267
QY 508 TCAAAATCAACACCGCGAATCTAGGCGATGGCACAGGACTAATAGAGAGATACCGGTG 567
Db 268 ProThrPheLeuThrProGluGlyAsnAlaSerAsnLysGluGluLeuArgGlyAsnVal 287
QY 568 TGGGGATCACCATGGATATATGAACGGAACCTCCGTCATTTGATAGTAAAGAAATG 627
Db 288 -----LeuSerLeuGluCysIleAlaGluGlyLeu 297
QY 628 GATGCCCGCTGTGTTTCCGTATTCATATTTGCA 663
Db 298 ProThr-----ProIleIleTyrTrpAla 305

```

RESULT 6
PCT-US03-28227-3411

```

; Sequence 3411, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELECEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 3411
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 930177.PT543p
PCT-US03-28227-3411

```

```

Alignment Scores:
Pred. No.: 2 94 Length: 1144
Score: 81.00 Matches: 42
Percent Similarity: 35.38% Conservative: 33
Best Local Similarity: 19.81% Mismatches: 77
Query Match: 6.75% Indels: 60
DB: 1 Gaps: 10

```

US-09-147-052-3_COPY_1_672 (1-672) x PCT-US03-28227-3411 (1-1144)

```

QY 67 CCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTCCGAGCGTCCAGTTGCTGAGGAA 126
Db 141 ProSerArgSerProLeuThrPheTrpThrLysGluLysLeuGluProIleThrLeuGlnSerGly 160
QY 127 GAGTCTACGTTTATCTTTGTCCTCCACAGGAGTGGTTCACCGTATCGGTGTAGAACCG 186
Db 161 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 175
QY 187 CCG-----CGAAATGTCGCGAACCTAGAAAGCCACC 219
Db 176 ProIleIlePheTrpMetAspAsnSerPheGlnArgLeuProGlnSerGluArgValSer 195
QY 220 GAGTGGGTGAAGGAATCCGATATATTAAAGAGAATATCAGTCCCATATAAATTTAA 279
Db 196 GlnGlyLeuAsnGly----- 200
QY 280 GTGAGCTTATATAAATATCAT---CAGACGACGACATGACGGGGGACGACATAT 336
Db 201 ---AspLeuTyPheSerAsnValLeuProGluAspThrArgGluAspTyrIleCystyr 219

```

```

QY 337 AGACAGATCACTAATCATATACA-----GATAGACGCCGCTTTCCATTGAGAGATC 390
Db 220 AlaArgPheAsnHisThrGlnGlnProLeuValLysValLysValLysValLysValLys 239
QY 391 ACGGATCTAATCGACGGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAAT 450
Db 240 Ser-----ValAsp-----GluLeuAsnAspThr 247
QY 451 GTATATGTTGAAGCTTTGACAGGGAT---GCGGGAGAAAAACAAGTACTTCTAAACCA 507
Db 248 IleAlaAlaAsnLeuSerAspThrGluPheTyGlyAlaLysSerArgGluArgPro 267
QY 508 TCAAAATTCACACGCCCGCAATAGGCGATGGCAGACACGATTAATGAGAGATACCGTG 567
Db 268 ProThrPheLeuThrProGluGlyAsnAlaSerAsnLysGluLeuArgGlyAsnVal 287
QY 568 TGGGGATCACCATTGATATATCGAACGGGAACCTCGTCAATGTATAGTAGAGGAAATG 627
Db 288 -----LeuSerLeuGluCysIleAlaGluGlyLeu 297
QY 628 GATGCCGCTCTGTGTTCCGTATTCATATTTGCA 663
Db 298 ProThr-----ProIleIleTyTrpAla 305

```

RESULT 7

```

PCT-US03-28227-3410
; Sequence 3410, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 3410
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 930177.PR537p
PCT-US03-28227-3410

```

Alignment Scores: 2.96 Length: 1179
Pred. No.: 81.00 Matches: 42
Score: 35.38% Conservative: 33
Percent Similarity:

Best Local Similarity: 19.81% Mismatches: 77
Query Match: 6.75% Indels: 60
DB: 1 Gaps: 10

US-09-147-052-3_COPY_1_672 (1-672) x PCT-US03-28227-3410 (1-1179)

```

QY 67 CCGAGTACCCAAATGTGACATCAAGAGAAGTGTTCGAGCGTCCAGTGTCTGAGGAA 126
Db 135 ProSerArgSerProLeuThrThrLysGluLysLeuGluProIleThrLeuGlnSerGly 154
QY 127 GAGTCTACCTTTTATCTTTGTCCTCCACAGGAGGGTTCACCGTGATCGCTAGAACCG 186
Db 155 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 169
QY 187 CGG-----CGAAAAATGTCGCCGAACCTAGAAAAAGCCACC 219
Db 170 ProIleIlePheThrMetAspAsnSerPheGlnArgLeuProGlnSerGluArgValSer 189
QY 220 GAGTGGGGTGAAGGAATCCGATATTTTAAAGAGAATATCAGTCCATATAAATTTAAA 279
Db 190 GlnGlyLeuAsnGly----- 194
QY 280 GTGACGCTTTATATAAAATATCATTT---CAGACGACGACATGAGCGGACGACATAT 336
Db 195 ---AspLeuTyPheSerAsnValLeuProGluAspThrArgGluAspTyrlleCysTyr 213
QY 337 AGACAGATCACAATCATATACA-----GATAGACGCCGCTTTCCATTGAGAGATC 390
Db 214 AlaArgPheAsnHisThrGlnThrIleGlnGlnLysGlnProIleSerValLysValLys 233
QY 391 ACGGATCTAATCGACGGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAAT 450
Db 234 Ser-----ValAsp-----GluLeuAsnAspThr 241
QY 451 GTATATGTTGAAGCTTTGACAGGGAT---GCGGGAGAAAAACAAGTACTTCTAAACCA 507
Db 242 IleAlaAlaAsnLeuSerAspThrGluPheTyGlyAlaLysSerArgGluArgPro 261
QY 508 TCAAAATTCACACGCCCGCAATAGGCGATGGCAGACACGACGATTAATGAGAGATACCGTG 567
Db 262 ProThrPheLeuThrProGluGlyAsnAlaSerAsnLysGluLeuArgGlyAsnVal 281
QY 568 TGGGGATCACCATTGATATATCGAACGGGAACCTCGTCAATGTATAGTAGAGGAAATG 627
Db 282 -----LeuSerLeuGluCysIleAlaGluGlyLeu 291
QY 628 GATGCCGCTCTGTGTTCCGTATTCATATTTGCA 663
Db 292 ProThr-----ProIleIleTyTrpAla 299

```

RESULT 8

```

PCT-US03-28227-5095
; Sequence 5095, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;

```

```
: APPLICANT: MA, Yan; JACKSON, Jennifer L.;
: APPLICANT: GIETZEN, Darryl; SUAREZ, Charlyn J.
: APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
: TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: PN-0100 PCT
: CURRENT APPLICATION NUMBER: PCT/US03/28227
: CURRENT FILING DATE: 2003-09-12
: PRIOR APPLICATION NUMBER: US 60/410,260
: PRIOR FILING DATE: 2002-09-12
: PRIOR APPLICATION NUMBER: US 60/410,259
: PRIOR FILING DATE: 2002-09-12
: NUMBER OF SEQ ID NOS: 5444
: SOFTWARE: PERL Program
: SEQ ID NO 5095
: LENGTH: 1224
: TYPE: PRT
: ORGANISM: Homo sapiens
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 930177.PT530P
PCT-US03-28227-5095

Alignment Scores:
Pred. No.: 2.98 Length: 1224
Score: 81.00 Matches: 42
Percent Similarity: 35.38% Conservative: 33
Best Local Similarity: 19.81% Mismatches: 77
Query Match: 6.75% Indels: 60
DB: 1 Gaps: 10

US-09-147-052-3_copy_1_672 (1-672) x PCT-US03-28227-5095 (1-1224)
QY 67 CCGAGTACCCAAATGTCACATCAAGAGAGTGTTCGAGCGTCCAGTTGCTGAGGAA 126
Db 141 ProSerArgSerProLeuThrPheThrLysGluLysLeuGluProIleThrLeuGlnSerGly 160
QY 127 GAGTCTACGTTTATCTTTGTCGCCACCGAGTGGTTCACCGGTGCTAGAACCG 186
Db 161 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 175
QY 187 CCG-----CGAAATGTCCCGAACCTAGAAAGCCACC 219
Db 176 ProIlePheThrMetAspSerPheGlnArgLeuProGlnSerGluArgValSer 195
QY 220 GAGTGGGTGAGCAATCGCATATTTAAAGAGATATCAGTCCATATAATTTAAA 279
Db 196 GlnGlyLeuAsnGly-----200
QY 280 GTGACGCTTTATATAAAATATCATT---CAGACGACGACATGGACGGGACGACATAT 336
Db 201 ---AspLeuThrPheSerAsnValLeuProGluAspThrArgGluAspThrIleCysTyr 219
QY 337 AGACAGATCACTAATCGATATACA-----GATAGGACGCCGCTTCCTCCATGAAGATC 390
Db 220 AlaArgPheAsnHisThrGlnThrIleGlnGlnLysGlnProIleSerValLysValle 239
QY 391 ACGGATCTATCGACGCAAGAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAAT 450
Db 240 Ser-----ValAsp-----GluLeuAsnAspThr 247
QY 451 GTATATGTTCAAGCGTTTGACAGGAT---GCGGGAGAAAACAAGTACTTCTAAACCA 507
Db 248 IleAlaValAsnLeuSerAspThrGluPheThrGlyAlaLysSerSerArgGluArgPro 267
QY 508 TCAAAATTCACACGCCGCAATCTAGGCGATGCGACACGACTAATGAGACGTATACCGTG 567
Db 268 ProThrPheLeuThrProGluGlyAsnAlaSerAsnLysGlnGluLeuArgGlyAsnVal 287
QY 568 TGGGGATCACCATGGATATTCGACACCGGAACTCCGTCATTTGATAGTAGAGAAATG 627
Db 288 -----LeuSerLeuGluCysIleAlaGluGlyLeu 297
QY 628 GATCGCGCTCTGTGTTCCGTATTCATATTTTGA 663
```

```
Db 298 ProThr-----ProIleLeuThrPala 305
RESULT 9
PCT-US03-28227-3408
: Sequence 3408, Application PC/TUS0328227
: GENERAL INFORMATION:
: APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
: APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
: APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
: APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
: APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
: APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
: APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
: APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
: APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
: APPLICANT: PANZER, Scott R.; WANG, Xinhao;
: APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
: APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
: APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
: APPLICANT: WU, Mingham C.; SPIRO, Peter A.;
: APPLICANT: LAGACE, Robert E.; WINGROVE, James A.;
: APPLICANT: STEWART, Elizabeth A.; KIRTON, Edward;
: APPLICANT: XU, Yuming; KWONG, Mary;
: APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
: APPLICANT: MA, Yan; JACKSON, Jennifer L.;
: APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
: APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
: TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: PN-0100 PCT
: CURRENT APPLICATION NUMBER: PCT/US03/28227
: CURRENT FILING DATE: 2003-09-12
: PRIOR APPLICATION NUMBER: US 60/410,260
: PRIOR FILING DATE: 2002-09-12
: PRIOR APPLICATION NUMBER: US 60/410,259
: PRIOR FILING DATE: 2002-09-12
: NUMBER OF SEQ ID NOS: 5444
: SOFTWARE: PERL Program
: SEQ ID NO 3408
: LENGTH: 1231
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 930177.PT528P
PCT-US03-28227-3408

Alignment Scores:
Pred. No.: 2.99 Length: 1231
Score: 81.00 Matches: 42
Percent Similarity: 35.38% Conservative: 33
Best Local Similarity: 19.81% Mismatches: 77
Query Match: 6.75% Indels: 60
DB: 1 Gaps: 10

US-09-147-052-3_copy_1_672 (1-672) x PCT-US03-28227-3408 (1-1231).
QY 67 CCGAGTACCCAAATGTCACATCAAGAGAGTGTTCGAGCGTCCAGTTGCTGAGGAA 126
Db 141 ProSerArgSerProLeuThrPheThrLysGluLysLeuGluProIleThrLeuGlnSerGly 160
QY 127 GAGTCTACGTTTATCTTTGTCGCCACCGAGTGGTTCACCGGTGCTAGAACCG 186
Db 161 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 175
QY 187 CCG-----CGAAATGTCCCGAACCTAGAAAGCCACC 219
Db 176 ProIlePheThrMetAspSerPheGlnArgLeuProGlnSerGluArgValSer 195
QY 220 GAGTGGGTGAGCAATCGCATATTTAAAGAGATATCAGTCCATATAATTTAAA 279
Db 196 GlnGlyLeuAsnGly-----200
```

RESULT 11
FCT-US03-28227-5093
Sequence 5093, Application PC/YUS0328227
: GENERAL INFORMATION:
: APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
: APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
: APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
: APPLICANT: HARTSHORNE, Tolnette A.; SUCHOROLSKI, Martin;
: APPLICANT: ALTUS, Christina M.; PITWS, Steven J.;
: APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
: APPLICANT: DEEGEANE, Angelo M.; PANESAR, Iqbal S.;
: APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
: APPLICANT: STEVENSON, Kristian A.; BLANCHARD, John L.;
: APPLICANT: PETERS, Scott R.; WANG, Xinhao;
: APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
: APPLICANT: PERALTA, Careyana H.; ANDERSON, Scott E.;
: APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
: APPLICANT: WU, Mingham C.; STUVE, Laura L.;

```
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5093
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 930177.PT526p
PCT-US03-28227-5093

Alignment Scores:
Pred. No.: 2,99 Length: 1234
Score: 81.00 Matches: 42
Percent Similarity: 35.38% Conservative: 33
Best Local Similarity: 19.81% Mismatches: 77
Query Match: 60 Indels: 10
DB: 1 Gaps: 1

US-09-147-052-3_COPY_1_672 (1-672) x PCT-US03-28227-5093 (1-1234)
QY 67 CCGAGTACCCAAATGTCATCAAGAGAGTTGTTTCGAGCGTCGAGTCTGAGGAA 126
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 141 ProSerArgSerProLeuThrThrLysGluLysLeuGluProIleThrLeuGlnSerGly 160
QY 127 GAGTCAGCTTTATCTTCCTCCCAACAGTGGTTCACCGTGATCCCTGACGACCG 186
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 161 GlnSerLeuValLeuProCysArgProIleGly-----LeuProPro 175
QY 187 CCG-----CGAAATGTCGCCAAGCTAGAAAGCCACC 219
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 ProIleIlePheThrMetAspAsnSerPheGlnArgLeuProGlnSerGluArgValSer 195
QY 220 GAGTGGGTGAGGAATCGCGATATTATTAAAGAGATATCATCGTCATATAATTTAA 279
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 GlnGlyLeuAsnGly-----GluLeuAsnAspThr 200
QY 280 GTGACGCTTTATATAAAATATCATTT---CAGACGACGACATGGAGCGGGGACGACATAT 336
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 201 ---AspLeuThrPheSerAsnValLeuProGluAspThrArgGluAspTyrIleCysTyr 219
QY 337 AGACAGATCATATCGATATACA-----GATAGGAGCGCGCTTCATCGATGAGAGATC 390
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 AlaArgPheAsnHisThrGlnThrIleGlnGlnLysGlnProIleSerValLysValIle 239
QY 391 ACGGATCTAATCGACGGCAAGAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAAT 450
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 Ser-----ValAsp-----GluLeuAsnAspThr 247
QY 451 GTATATGTTGAAGCGTTTTCACAGGGAT---CGGGGAGAAAAACAAGTACTTCTAAAACCA 507
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 248 IleAlaAlaAsnLeuSerAspThrGluPheThrGlyAlaLysSerSerArgGluArgPro 267
QY 508 TCAAAATTCACACGCCCGAATCTAGGCGATGAGCGACGACGACTAATGAGACGATACCGGTG 567
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 268 ProThrPheLeuThrProGluGlyAsnAlaSerAsnLysGluGluLeuArgGlyAsnVal 287
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 5093
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 930177.PT526p
PCT-US03-28227-5093

Alignment Scores:
Pred. No.: 2,88 Length: 385
Score: 80.00 Matches: 38
Percent Similarity: 29.94% Conservative: 12
Best Local Similarity: 22.75% Mismatches: 49
Query Match: 6 Indels: 6
DB: 7 Gaps: 7

US-09-147-052-3_COPY_1_672 (1-672) x US-10-389-647-695 (1-385)
QY 301 ATCATTTACAGACGACATGAGGGGACGACACA-----TATAGACAGATCATCTAATCA 354
Db : : : : : ||||| : : : : : : : : : : : : : : : : : : :
QY 201 ValPheSerThrValValGlyThrGlyThrAlaGluLeuThrArgAlaIleAlaArgArg 220
QY 355 TATACAGATAGGACGCCCGTTTCCATTGAGAGATCAGCGATCTATCCGCGCAAGCA 414
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 Tyr-----GlyAspGly 224
QY 415 AGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGACAGG 474
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 ArgArgProIleAlaSerLeuThrThrSerGluAlaGluValAlaLysMetGluSer 244
QY 475 GATCGCGGGAGAAAACAAGTACTTCTTAACCA-----TCAAAATTCACACGCCCGGAA 528
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 AspValAlaGluGlyGlnValValAlaProTyrPheSerSerIleAspThrAlaAla 264
QY 529 TCTAGGCA-----TGCCACACGACTAATGAGAGCTATACCGGTG 567
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 SerArgAlaPheValGlnAlaCysHisGlyPhePheProGluAsnAlaThrIleThrAla 284
QY 568 TGGGGATCACCATTGATATATCGAACG-----GGAACCTCCGTCAT----- 609
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 TrpAlaGluAlaLysTrpGlnThrLeuLeuGlyArgAlaAlaGlnAlaAlaGly 304
QY 610 -----TGTATA----- 615
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 305 SerTrpArgValGluAspValGlnArgHisLeuTyrAspIleCysIleAspAlaProGln 324
QY 616 -----GTAGAGAA----- 624
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 GlyProValArgValGluArgGlnAsnAsnHisSerArgLeuSerSerArgIleAlaGlu 344
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 625 ATGGATGCCGCTCTGTGTTT 645
Db 345 ILeaspAlaAa-ggLyValPhe 351

RESULT 13

US-10-425-114A-58466
; Sequence 58466, Application US/10425114A
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114A

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 58466

; LENGTH: 947

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700383178_FLI.pep

US-10-425-114A-58466

Alignment Scores:
Pred. No.: 5,72 Length: 947
Score: 78.00 Matches: 45
Percent Similarity: 38.96% Conservative: 45
Best Local Similarity: 19.48% Mismatches: 91
Query Match: 6.50% Indels: 50
DB: 6 Gaps: 10

US-09-147-052-3_COPY_1_672 (1-672) x US-10-425-114A-58466 (1-947)

QY 55 ACGAACTCATCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAG 114
Db 625 SerAsnValAsnProSerGlyLysGlnCysProAspAsnSerAlaSerCysLysThrAla 644
QY 115 TTGCTGAGGAGAGTCTACGTTTATCTTTGT-----CCC 150
Db 645 LeuMetAspValProSerThrSerHisLeuCysSerIleAspLeuLysArgCysGluSer 664
QY 151 CCACAGTGGTTCACACCGTGTAGAACCGCGCGAAATGTCGGAACCTAGA 210
Db 665 ProSerValProSerLeuLeuIleLysHisGluLeuAlaGlnPheAlaLysGlnProCys 684
QY 211 AAGCCACCGAGTGGGGTGAAGGAATCCGGATATATTATTAAGAGAATATCAGTCCATAT 270
Db 685 LeuValAspLysSerGluAspGly-----CysLeuSerGlnAsnAsnGlyMethHis 701
QY 271 AAATTTAAAGTCACGCTTTATATAAAAT-----ATCATTTCAG 309
Db 702 LysSerGlnIleArgLeuTrpIleGluSerGlyGlnAsnMetLysCysGluSerAlaSer 721
QY 310 ACGACGACATGACGCGGACACATATAGACAGATCACTAATCGATATACAGATAGGACG 369
Db 722 AlaSerAsnSerLysGlyLeuAspValProSerLysAlaAsnArgLysArgSerArgPro 741
QY 370 CCCGTTTCC-----ATTGAAGAG 387
Db 742 GlyGluSerProLysProArgProLysAspArgGlnLeuIleGlnAspArgIleLysGlu 761
QY 388 ATCAGGATCTAATCGAGGCAAGAGATGCTCATCTAAGCA-----AGATAC 438
Db 762 LeuArgGluLeuValProAsnGlyAlaLysCysSerIleAspGlyLeuLeuGluLysThr 781
QY 439 CTTAGAAACATGTATATGTTCAAGCGTTTGACACGGATCGGGGAGAAAACAACTACTT 498
Db 782 ValLysHisMetLeuPheLeuGlnSerValThrLysAsnAlaAspLys----- 797

QY 499 CTAATAACCATCAAAATTCACACGCCGAATCTAGGCGATGGCACACAGACTAAT----- 552
Db 798 LeuLysAspSer-----ThrGluSerLysIleLeuGlySerGluAsnGlyPro 813
QY 553 -----GAGACGTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACC---TCC 603
Db 814 LeuTrpLysAspTyrPheGluGlyAlaThrPheAlaPheAspValGlySerGlnSer 833
QY 604 GTCATTTGT-----ATAGTAGAGAAATGGAT 630
Db 834 MetThrCysProIleIleValGluAspLeuAsp 844

RESULT 14

US-10-425-114A-70492

; Sequence 70492, Application US/10425114A

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114A

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 70492

; LENGTH: 965

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLB73025E08_FLI.pep

US-10-425-114A-70492

Alignment Scores:
Pred. No.: 5,74 Length: 965
Score: 78.00 Matches: 45
Percent Similarity: 38.96% Conservative: 45
Best Local Similarity: 19.48% Mismatches: 91
Query Match: 6.50% Indels: 50
DB: 6 Gaps: 10

US-09-147-052-3_COPY_1_672 (1-672) x US-10-425-114A-70492 (1-965)

QY 55 ACGAACTCATCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAG 114
Db 643 SerAsnValAsnProSerGlyLysGlnCysProAspAsnSerAlaSerCysLysThrAla 662
QY 115 TTGCTGAGGAGAGTCTACGTTTATCTTTGT-----CCC 150
Db 663 LeuMetAspValProSerThrSerHisLeuCysSerIleAspLeuLysArgCysGluSer 682
QY 151 CCACAGTGGTTCACACCGTGTAGAACCGCGCGAAATGTCGGAACCTAGA 210
Db 683 ProSerValProSerLeuLeuIleLysHisGluLeuAlaGlnPheAlaLysGlnProCys 702
QY 211 AAGCCACCGAGTGGGGTGAAGGAATCCGGATATATTATTAAGAGAATATCAGTCCATAT 270
Db 703 LeuValAspLysSerGluAspGly-----CysLeuSerGlnAsnAsnGlyMethHis 719
QY 271 AAATTTAAAGTCACGCTTTATATAAAAT-----ATCATTTCAG 309
Db 720 LysSerGlnIleArgLeuTrpIleGluSerGlyGlnAsnMetLysCysGluSerAlaSer 739
QY 310 ACGACGACATGACGCGGACACATATAGACAGATCACTAATCGATATACAGATAGGACG 369
Db 740 AlaSerAsnSerLysGlyLeuAspValProSerLysAlaAsnArgLysArgSerArgPro 759
QY 370 CCCGTTTCC-----ATTGAAGAG 387

Db 760 GlyLysProLysProArgProLysAspArgGlnLeuIleLeuAspArgIleLysGlu 779
QY 388 ATCAGCGATCTATCATCGCGGAAAGAGAGATCTCTAAGCA-----AGATAC 438
Db 780 LeuArgGluLeuValProAsnGlyAlaLysCysSerIleAspGlyLeuLysGluLysThr 799
QY 439 CTAGAGAAACAATGTATATGTTGAAGCGTTTGACAGGGATCGCGGAGAGAAACAAGTACTT 498
Db 800 ValLysHisMetLeuPheLeuGlnSerValThrLysAsnAlaAspLys-----815
QY 499 CTAAGACCAATCAAAATCAACAGCGCCGGAATCTAGGGCATGGCAGCAGCTAAT-----552
Db 816 LeuLysAspSer-----ThrGluSerLysIleLeuGlySerGluAsnGlyPro 831
QY 553 -----GAGACGTATACCGTGTGGGATCACCATGATATATCGAAGCGGAACC---TCC 603
Db 832 LeuTrpLysAspTrpPheGluGlyGlyAlaThrTrpAlaPheAspValGlySerGlnSer 851
QY 604 GTCAATGT-----ATAGTAGAGAAATGGAT 630
Db 852 MetThrCysProIleIleValGluAspLeuAsp 862

RESULT 15
US-10-425-114A-71438
; Sequence 71438, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71438
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMOL7041C02_FLI.pep
US-10-425-114A-71438

Alignment Scores:
Pred. No.: 6.35 Length: 541
Score: 77.00 Matches: 35
Percent Similarity: 37.58% Conservative: 21
Best Local Similarity: 23.49% Mismatches: 63
Query Match: 6.42% Indels: 30
DB: 6 Gaps: 6

US-09-147-052-3_COPY_1_672 (1-672) x US-10-425-114A-71438 (1-541)
QY 208 AGAAAGCCAGCGAGTGGGTGAGGAGATCCGGATATATTAAAGAGATATCAGTCCA 267
Db 90 LysArgAlaAsnGluArgGlyGluThrValThrSerLeuSerSerGlnPheIleAsnGlu 109
QY 268 TATAAATTAAAGTCAGCGCTTTATTATAAAATATCATTCAGACGACATGGACGGG 327
Db 110 PheLeuLeuAspMetThr-----GluLeuGlnCysLeuProThr----123
QY 328 ACGACATATAGACAGATCACTAATCGATATACATAGACGCGCGGTTTCCATTGAAGAG 387
Db 124 -----CysGluProArgValThrGluHisIleGluHisIleLysLeu 138
QY 388 ATCAGCGATCTAATCGCGGAAAGAGAGATCTCTAAGCAAGAGATACCTTAGAAC 447
Db 139 IleThrGlnIleMetAsp-----AsnGlyLysAlaIleGluGly 153
QY 448 AATGTATAT-----GTTGAGCG-----TTT 468

Db 154 AspValTyPheSerValGluSerPheProGluTyTrpLeuSerLeuSerGlyArgLysPhe 173
QY 469 GACAGGGATCGCGGAGAAAAACAAGTACTTCTAAAAACCATCAAAATTCACACACCGCGAA 528
Db 174 AspGlnAsnGlnAlaGlyAlaArgValAlaPheAspThrArgLysArgAsnProAlaAsp 193
QY 529 TCTAGGGCATGGCAGACACGACTAATAGACAGCTATACCGTGTGGGATCACCATGG---ATA 585
Db 194 PheAlaLeuTrpLysAlaAlaLysGluGlyGluProPheTrpAspSerProTrpGlyArg 213
QY 586 TATCGAACGGGAACCTCCGTCATGT 612
Db 214 GlyArgProGlyTrpHisIleGluCys 222

Search completed: October 8, 2003, 19:18:12
Job time : 16.5126 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:53:53 ; Search time 203.34 Seconds
(without alignments)
6014.205 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200
Sequence: 1 atgcacatttttagcgga.....catattttgcaatggccaat 672

Scoring table:

BLOSUM62 Xgapop 10.0, Xgapext 0.5
Xgapop 10.0, Ygapext 0.5
Xgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 11457514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=slp
-Q=/cgn2_1/USPTO.spool/US09147052/runat_08102003_154343_29901/app_query.fasta_1.5980
-DB=Pending_Patents_AA_Main -OPMT=fastan -SUFFIX=rapm -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052_@CGN_1_1_2602_@runat_08102003_154343_29901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	100.0	865	3	US-07-722-860-13
2	1200	100.0	865	6	US-08-213-449A-13
3	1200	100.0	865	6	US-08-213-449B-13
4	1200	100.0	1086	15	US-09-147-052-4
5	606	50.5	868	12	US-08-804-439-21
6	606	50.5	868	17	US-09-301-390-21
7	606	50.5	868	17	US-09-338-326-21
8	583	48.6	943	4	US-08-096-183D-4
9	583	48.6	943	19	US-09-521-738-14
10	582	48.5	933	3	US-07-805-524-2
11	513	42.8	891	9	US-08-541-878-6
12	504	42.0	885	9	US-08-541-878-8
13	504	42.0	885	12	US-08-804-439-23
14	504	42.0	885	17	US-09-301-390-23
15	504	42.0	885	17	US-09-338-326-23
16	504	42.0	903	12	US-08-804-439-22
17	504	42.0	903	17	US-09-301-390-22
18	504	42.0	903	17	US-09-338-326-22
19	504	42.0	904	32	US-60-412-956-12
20	498	41.5	904	1	PCT-US03-11231-18
21	498	41.5	904	27	US-10-121-988-18
22	498	41.5	904	28	US-10-200-562-18
23	498	41.5	904	28	US-10-237-551-18
24	482	40.2	846	5	US-08-123-456-142
25	482	40.2	846	16	US-09-297-477A-142
26	482	40.2	846	25	US-09-994-404-142
27	482	40.2	896	5	US-08-123-456-201
28	482	40.2	896	16	US-09-297-477A-241
29	482	40.2	896	25	US-09-994-404-241
30	454.5	37.9	250	19	US-09-521-738-16
31	453	37.8	787	5	US-08-123-456-79
32	453	37.8	787	16	US-09-297-477A-79
33	453	37.8	787	25	US-09-994-404-79
34	314	26.2	62	27	US-10-131-591A-5
35	314	26.2	456	15	US-09-147-052-2
36	314	26.2	456	24	US-09-301-572A-3
37	302	25.2	62	27	US-10-131-591A-6
38	235	19.6	200	22	US-09-791-537-145956
39	225	18.8	823	26	US-10-055-364-37
40	225	18.8	845	12	US-08-804-439-94
41	225	18.8	845	17	US-09-301-390-94
42	225	18.8	845	17	US-09-338-326-94
43	219	18.2	844	26	US-10-055-364-40
44	219	18.2	874	12	US-08-804-439-15
45	219	18.2	874	17	US-09-301-390-15

ALIGNMENTS

RESULT 1

US-07-722-860-13
; Sequence 13, Application US/07722860
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohel
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR

TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 STREET: 301 North Washington Street
 CITY: Falls Church
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: US/07/722,860
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Murphy Jr., Gerald M.
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 1644-101P
 TELEPHONE: (703) 241-1300
 TELEFAX: (703) 241-2848
 TELEX: 248345
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 865 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-722-860-13

Alignment Scores:
 Pred. No.: 4.62e-124 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-07-722-860-13 (1-865)

QY 1 ATGCACATATTTAGCGGAATGTCATATTTTCCTATAGTTATCTATATGTCACGAC 60
 Db 1 MethistyrPheArgArGAsnCysilePhePheLeuileValileLeutyrglyThrAsn 20
 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAACTTGTTCGAGCGTCCAGTTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
 QY 121 GAGGAAGTCTACGTTTATCTTTGTCCCCACCGAGTGGTTCACCGGTGATCCGTCTA 180
 Db 41 GluGluGluSerThrPheThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 QY 181 GAACCCCGCGAAATGTCGACACCTAGAAAGCCAGTGGGGTCAAGGAGTCGG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGlyLeuLeu 80
 QY 241 ATATATTTAAAGAAATATCATGTCATATAATTTAAAGTACGCTTTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProThrLysPheLysValThrLeuTyTyLysAsn 100
 QY 304 ATCATTCAGACGACATGTCAGCGGGACGACATATAGACATCACTAATCGATATACA 360
 Db 101 IleTyrGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
 QY 361 GATAGGACGCGCTTTCATTGAGACATCAGGATCAATCGACGCAAGGAGATGC 420
 Db 121 AspArgThrProValSerIleGluIleThrAspLeuileAspGlyLysGlyArgCys 140
 QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAACGGTTTGACAGGGATCG 480

Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla 160
 QY 481 GGAAAAACAAGTACTTCTTAAACACCATCAAAATTCACACGCCCGAATCTAGGCGATGG 540
 Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
 QY 541 CACACGACTAATGACAGCGTATACCGTGTGGGATCACCATGATATATATCAACGGGAAC 600
 Db 181 HistThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
 QY 601 TCCGTCATTTGTATAGTAGAGAAATGGATGCCCTCTGTGTTCCGTATTCATATTTT 660
 Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSeryrPhe 220
 QY 661 GCAATGGCAAT 672
 Db 221 AlaMetAlaAsn 224

RESULT 2
 US-08-213-449A-13
 ; Sequence 13, Application US/08213449A
 ; GENERAL INFORMATION:
 ; APPLICANT: NAZERIAN, Keyvan
 ; APPLICANT: LEE, Lucy F.
 ; APPLICANT: YANAGIDA, Noboru
 ; APPLICANT: OGAWA, Ryohel
 ; APPLICANT: LI, Yi
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
 ; PROTECTION AGAINST MAREK'S DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 15-MAR-1994
 ; APPLICATION NUMBER: US/08/213,449A
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M.
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 1644-108P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 865 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-213-449A-13

Alignment Scores:
 Pred. No.: 4.62e-124 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-08-213-449A-13 (1-865)

QY 1 ATGCACATATTTAGCGGAATGTCATATTTTCCTATAGTTATCTATATGTCACGAC 60

Db 221 AlaMetaLaasn 224

RESULT 4

US-09-147-052-4

; Sequence 4, Application US/09147052

; GENERAL INFORMATION:

; APPLICANT: SAITOH, Shuji

; APPLICANT: TSUZAKI, Yoshinari

; APPLICANT: YANGIDA, Noboru

; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,

; FILE REFERENCE: 981167

; CURRENT FILING DATE: 1999-04-05

; PRIOR FILING DATE: 1996-03-29

; PRIOR FILING DATE: 1997-03-28

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1086

; TYPE: PRT

; ORGANISM: hybrid

US-09-147-052-4

Alignment Scores:

Pred. No.: 4,99e-124 Length: 1086

Score: 1200.00 Matches: 224

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 15 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-09-147-052-4 (1-1086)

QY 1 ATGCACTATTTAGCGGAATTCATATTTTCCTTATAGTATCTATATCGTACGAAC 60

Db 1 MethIstYrPheAgaGAsnCysIlePheHeLeuIleValIleLeuTyrGlyThrAsn 20

QY 61 TCATCTCGAGTACCCAAATGTGACATCAAGAGAAGTGTTCGAGCGTCCAGTGTGCT 120

Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValSerValGlnLeuSer 40

QY 121 GAGGAAGAGTCTAGTTTATCTTTGTCCTCCACAGTGGTTCACCGTGATCCGCTA 180

Db 41 GluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60

QY 181 GAACCGCCGCGAAATGTCCTCGAACCTAGAAAGCCACCGAGTGGGTGAAGAAATCGG 240

Db 61 GluProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80

QY 241 ATATTATTAAGAGATATACATCCATATAATTTAAAGTGACGCTTTATTATTAAT 300

Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100

QY 301 ATCATTCAGACGACATGAGCGGGGAGACATATAGACATCACTAATCATATACA 360

Db 101 IleIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120

QY 361 GATAGACCGCGTTCATTCATTCAGAGATACAGGATCAATACGCGGCAAGAGATGC 420

Db 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140

QY 421 TCATCTAAGCAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTGACAGGATCG 480

Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla 160

QY 481 GGAGAAAACAAGTACTCTAAACCCATCAAAATTCACACGCCCGCATCTAGGCGATGG 540

Db 161 GlyGlyLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180

QY 541 CACAGCACTAATGAGCGTATACCGTGTGGGGATCACCATGATATATCGAACGGGAAC 600

Db 181 HlsThrAsnGluThrThrValTrpGlySerProTyrPilleTyrArgThrGlyThr 200

QY 601 TCCGTCAATTGTATAGTAGAGAAATGATGATCCCTCTGCTGTTCCGTATTCATATTTT 660

Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220

QY 661 GCAATGGCCAAT 672

Db 221 AlaMetaLaasn 224

RESULT 5

US-08-804-439-21

; Sequence 21, Application US/08804439

; GENERAL INFORMATION:

; APPLICANT: Rose, Timothy M.

; APPLICANT: Bosch, Marnix L.

; APPLICANT: Strand, Kurt

; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV

; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES

; NUMBER OF SEQUENCES: 113

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster

; STREET: 755 Page Mill Road

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/804,439

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Schiff, J. Michael

; REGISTRATION NUMBER: 40,253

; REFERENCE/DOCKET NUMBER: 29938-20002.20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 813-5600

; TELEFAX: (415) 494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 668 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-804-439-21

Alignment Scores:

Pred. No.: 9.47e-58 Length: 868

Score: 606.00 Matches: 107

Percent Similarity: 74.09% Conservative: 36

Best Local Similarity: 55.44% Mismatches: 50

Query Match: 50.50% Indels: 0

DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-08-804-439-21 (1-868)

QY 91 AGAGAAGTGTTCGAGCGTCAGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTC 150

Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60

QY 151 CCACGAGTGGTTCACCGTGTATCGCTAGAACCGCGCGAAATGTCGCGAACCTAGA 210

Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80

QY 211 AAAGCCACGAGTGGGTGAAGGAATCCGATATATTTAAAGAGATATATCAGTCCATAT 270

```

Db      81 LeuGlyLysAsnProThrGluGlyIleAlaValValTyrLysGluAsnIleAlaAlaTyr 100
      ::: ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      271 AAATTAAAGTCAGCGCTTTATATAAAATATCATTCAGACGACATGGACGGGACG 330
      ||||||| ||::: ||||||| :::::::::::::: ||||||| ||| ||:::
Db      101 LysPheLysAlaThrValTyrLysAspValIleValSerThrAlaTrpAlaGlySer 120
      ||||||| ||::: ||||||| :::::::::::::: ||||||| ||| ||:::
QY      331 ACATATAGACAGATCACTAATCATATACAGATAGGACCGCGTTTCCATTGAAGAGATC 390
      :::: ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      391 ACGGATCTAATCGCGGCAAGGAGATCGTCATCAAGCAAGATACCTTAGAACCAAT 450
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      141 ThrAspThrIleAspLysPheGlyLysCysSerLysAlaThrTyrValArgAsnAsn 160
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      451 GTATATGTGAAGCGTTTGACAGGATCGCGGAGAAAACAAGTACTTCTAAACCATCA 510
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      511 AAATTCACACGCCCGAATCTAGGCGATGGCACAGCACTAATGACACGTATACCGTGTGG 570
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      181 LysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnAspThrTyrMetValAla 200
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      571 GGATCACCATGATATATCGAAGCGGAACCTCCGTCATTTGTATAGTAGAGAAATGCAT 630
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      631 GCCCGCTCTGTTTCCGTTATTCATATTTTGCATGGCC 669
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSer 233
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||

```

RESULT 6

```

US-09-301-390-21
; Sequence 21, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Mark L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-301-390-21
Alignment Scores:
Pred. No.: 9,47e-58 Length: 868
Score: 606.00 Matches: 107
Percent Similarity: 74.09% Conservative: 36
Best Local Similarity: 55.44% Mismatches: 50
Query Match: 50.50% Indels: 0
DB: 17 Gaps: 0
US-09-147-052-3_COPY_1_672 (1-672) x US-09-301-390-21 (1-868)
QY      91 AGAGAAGTGTTCCTGAGCGTCCAGTTGTCTGAGGAGAGTCTACGTTTATCTTGTGTC 150
      ||||||| :::: ||| ||::: ||||||| :::: ||||||| :::: |||
Db      41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
      ||||||| :::: ||| ||::: ||||||| :::: ||||||| :::: |||
QY      151 CCACGAGTGGTTCACCGTGATCCGTGTAGAACCGCCGCGAAATGTCCCGAACCTAGA 210
      ||||||| ||||||| :::: ||| ||::: ||||||| :::: |||
Db      61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      211 AAAGCCACCGAGTGGGTGAAGGAATCGCGATATATTATTAAAGAGAATATCATGTCATAT 270
      ||||||| :::: ||| ||::: ||||||| :::: ||||||| :::: |||
Db      81 LeuGlyLysAsnProThrGluGlyIleAlaValValTyrLysGluAsnIleAlaTyr 100
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      271 AAATTAAAGTCAGCGCTTTATATAAAATATCATTCAGACGACGACATGGACGGGACG 330
      ||||||| ||::: ||||||| :::::::::::::: ||||||| ||| ||:::
Db      101 LysPheLysAlaThrValTyrLysAspValIleValSerThrAlaTrpAlaGlySer 120
      ||||||| ||::: ||||||| :::::::::::::: ||||||| ||| ||:::
QY      331 ACATATAGACAGATCACTAATCATATACAGATAGGACCGCGTTTCCATTGAAGAGATC 390
      :::: ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      391 ACGGATCTAATCGCGGCAAGGAGATCGTCATCTAAGCAAGATACCTTAGAACCAAT 450
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      141 ThrAspThrIleAspLysPheGlyLysCysSerLysAlaTrpHisThrThrAsnAspThrTyrMetValAla 200
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      451 GTATATGTGAAGCGTTTGACAGGATCGCGGAGAAAACAAGTACTTCTAAACCATCA 510
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      511 AAATTCACACGCCCGAATCTAGGCGATGGCACAGCACTAATGACACGTATACCGTGTGG 570
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      181 LysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnAspThrTyrMetValAla 200
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      571 GGATCACCATGATATATCGAAGCGGAACCTCCGTCATTTGTATAGTAGAGAAATGCAT 630
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
QY      631 GCCCGCTCTGTTTCCGTTATTCATATTTTGCATGGCC 669
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||
Db      221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSer 233
      ||||||| ||||||| ||||||| ||||||| :::::::::::::: ||||||| :::: |||

```

RESULT 7

```

US-09-338-326-21
; Sequence 21, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Mark L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

QY 7 TATTTAGCGGAATTCATATTT----- 30
Db 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCCTATAT-----GGTACGAATCATCTCCG 69
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 70 AGTACCCAAATGTGCATCAAGAGAAGTT----- 99
Db 82 ProArgArgThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
QY 100 -----GTTTCGAGCGTCCAGTGTCTCGAGGAAGAG----- 129
Db 102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
QY 130 TCTACGTTTATCTTTGTCCTCCACAGTGGTTCACCGGTCAACCGGTATCGTCTAGAACCGCG 189
Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
QY 190 CGAAATGTCGCGAACCTAGAAAGCCACCGAGTGGGTGAAGGAATCGCATATATTT 249
Db 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
QY 250 AAAGAGATATCAGTCCATATAAATTTAAAGTGCAGCTTTTATATAAATATCATTCAG 309
Db 162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
QY 310 ACGAGCATATGACGGGACGACATATAGACAGATCACTAATCGATATACAGATAGGACG 369
Db 182 ThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201
QY 370 CCCGTTTCATAGAGATCAGCATCTAATCGAGCGCAAGGAGAGATGCTCACTATAA 429
Db 202 ProValLysValGlnIleThrAspLeuIleAspArgArgGlyMetCysLeuSerLys 221
QY 430 GCAAGTACTCTAAACCACTAAATTCACACCGCCGAATCTAGGCGATGGCACACGACT 549
Db 242 GluLeuProLeuLysProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260
QY 550 AATGAGAGTATACCGTGTGGGATCACCATGGATATATATCGAACGGACCTCCGTCAT 609
Db 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn 280
QY 610 TGTATAGTAGAGAAATGGATCCCGCTGTGTTCGTTCCGTATCATATTTTCCATGGCC 669
Db 281 CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300

RESULT 9
US-09-521-738-14
; Sequence 14, Application US/09521738
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9
; CURRENT APPLICATION NUMBER: US/09/521,738
; CURRENT FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 943

```

```

; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
; US-09-521-738-14

Alignment Scores:
Pred. No.: 3,6e-55 Length: 943
Score: 583.00 Matches: 120
Percent Similarity: 56.07% Conservative: 37
Best local Similarity: 42.86% Mismatches: 63
Query Match: 48.58% Indels: 60
DB: 19 Gaps: 5

US-09-147-052-3_COPY_1_672 (1-672) x US-09-521-738-14 (1-943)

QY 7 TATTTAGCGGAATTCATATTT----- 30
Db 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCCTATAT-----GGTACGAATCATCTCCG 69
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 70 AGTACCCAAATGTGCATCAAGAGAAGTT----- 99
Db 82 ProArgArgThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
QY 100 -----GTTTCGAGCGTCCAGTGTCTCGAGGAAGAG----- 129
Db 102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
QY 130 TCTACGTTTATCTTTGTCCTCCACAGTGGTTCACCGGTCAACCGGTATCGTCTAGAACCGCG 189
Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
QY 190 CGAAATGTCGCGAACCTAGAAAGCCACCGAGTGGGTGAAGGAATCGCATATATTT 249
Db 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
QY 250 AAAGAGATATCAGTCCATATAAATTTAAAGTGCAGCTTTTATATAAATATCATTCAG 309
Db 162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
QY 310 ACGAGCATATGACGGGACGACATATAGACAGATCACTAATCGATATACAGATAGGACG 369
Db 182 ThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201
QY 370 CCCGTTTCATAGAGATCAGCATCTAATCGAGCGCAAGGAGAGATGCTCACTATAA 429
Db 202 ProValLysValGlnIleThrAspLeuIleAspArgArgGlyMetCysLeuSerLys 221
QY 430 GCAAGTACTCTAAACCACTAAATTCACACCGCCGAATCTAGGCGATGGCACACGACT 549
Db 242 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241
QY 490 CAAGTACTCTAAACCACTAAATTCACACCGCCGAATCTAGGCGATGGCACACGACT 549
Db 242 GluLeuProLeuLysProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260
QY 550 AATGAGAGTATACCGTGTGGGATCACCATGGATATATATCGAACGGACCTCCGTCAT 609
Db 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn 280
QY 610 TGTATAGTAGAGAAATGGATCCCGCTGTGTTCGTTCCGTATCATATTTTCCATGGCC 669
Db 281 CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300

RESULT 10
US-07-805-524-2
; Sequence 2, Application US/07805524

```

GENERAL INFORMATION:
 APPLICANT: Babluk, Lorne
 APPLICANT: Van Der Hurk, Sylvia
 APPLICANT: Zamb, Tim
 APPLICANT: Fitzpatrick, David
 TITLE OF INVENTION: Bovine Herpesvirus type 1 Polypeptides
 TITLE OF INVENTION: and Vaccines
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1018

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/805,524
 FILING DATE: 19911211
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Gracey, Nancy J.
 REGISTRATION NUMBER: 28,216
 REFERENCE/DOCKET NUMBER: 293102000520
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-813-5600
 TELEFAX: 415-494-0792
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 933 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-805-524-2

Alignment Scores:
 Pred. No.: 4.64e-55
 Score: 582.00
 Percent Similarity: 68.45%
 Best Local Similarity: 51.94%
 Query Match: 48.50%
 DB: 3
 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-07-805-524-2 (1-933)

```

QY 52 GGTACGAACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTCCAGCGTC 111
Db 97 GlyAspAlaAlaSerProAspAsnSerThrAspValargAlaAlaLeuArgLeuAla 116
QY 112 CAGTTGCTCAGAGAAGTCTAGTTTATCTTTGTCCTCCACCAGTGGTTCAACCGTG 171
Db 117 GlnAlaAlaGlyGluAsnSerArgPheValCysProProSerGlyAlaThrVal 136
QY 172 ATCCGCTAGAACCGCGGCGAAATGTCGGRACCTAGAAACCCAGCGAGTGGGTAA 231
Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
QY 232 GGATCCGATATATTTAAAGAGAATATCAGTCCATATAAAATTTAAAGTACGCTTAT 291
Db 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIleTyr 176
QY 292 TATAAATATCATTCAGACGACGATCGCGGAGCAGCATATATACAGATCACTAAT 351
Db 177 TyrLysAsnValIleValThrThrThrTrpAlaGlySerThrTyrAlaAlaIleThrAsn 196
QY 352 CCATATACAGATAGGACCGCGCTTCCATTAAGAGATCAGGATCTAATCGACGCAAA 411
Db 197 GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLysLys 216

```

```

QY 412 GGAAGATGCTCATCTAAGCAAGATACCTTAGAACCAATGTATATGTGAGCGTTTGAC 471
Db 217 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValValAlaPheasp 236
QY 472 AGGATGCGGAGAGAAACAAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCT 531
Db 237 ArgAspAspProTrpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 256
QY 532 AGGCGATGCGACAGCACTAATGAGACGTATACCGTGTGGGATCACCATGATATATCGA 591
Db 257 ArgGlyTrpHisThrThrAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 276
QY 592 ACGGGAACCTCCGTCATCTATAGTAGAGGAATGATCCCGCTCTGTGTTTCCGTAT 651
Db 277 ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 296
QY 652 TCATATTTTGCATGGCC 669
Db 297 AspSerPheAlaLeuSer 302

```

RESULT 11
 US-08-541-878-6
 ; Sequence 6, Application US/08541878
 ; GENERAL INFORMATION:
 ; APPLICANT: Eberle, Richard
 ; APPLICANT: Black, Darla
 ; APPLICANT: Scinicariello, Franco
 ; APPLICANT: Hilliard, Julia K.
 ; TITLE OF INVENTION: Cloning and Amplification of Monkey B
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cox & Smith Incorporated
 ; STREET: 112 East Pecan Street, Suite 2000
 ; CITY: San Antonio
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 78205
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/541,878
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/042,747
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haymond, W. Bradley
 ; REGISTRATION NUMBER: 35186
 ; REFERENCE/DOCKET NUMBER: S-0072.179
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 210-554-5500
 ; TELEFAX: 210-226-8395
 ; TELEX: 767609
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 891 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-541-878-6

Alignment Scores:
 Pred. No.: 2.3e-47
 Score: 513.00
 Percent Similarity: 63.59%
 Best Local Similarity: 46.12%
 Query Match: 42.75%
 Length: 891
 Matches: 95
 Conservative: 36
 Mismatches: 75
 Indels: 0


```

Db      268 PheValLeuAla 271
RESULT 13
US-08-804-439-23
; Sequence 23, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439-23
Alignment Scores:
Pred. NO.: 2,32e-46 Length: 885
Score: 504.00 Matches: 94
Percent Similarity: 62.75% Conservative: 34
Best Local Similarity: 46.08% Mismatches: 76
Query Match: 42.00% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-08-804-439-23 (1-885)
QY      58 AACTCATCTCCGATCCCAAAATGTGACATCAAGAGAAGTTGTTTCGACGCGTCCAGTTG 117
      |||:|||||
Db      68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleLysAla 87
      |||:|||||
QY      118 TCTGAGGAAGACTAGTCTTTATCTTGTGCCCCCACCAGTGGTTCAACCGGTATCCGT 177
      :||:|||||
Db      88 ArgAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValGln 107
      :||:|||||
QY      178 CTAGAACCCCGCGAAATGTCCCGAACCTAGAAAAAGCCACCGAGTGGGTGAAGAAATC 237
      |||:|||||
Db      108 PheGluGlnProArgProCysProArgAlaProAspGlyGlnAsnTyrThrGluGlyIle 127
      |||:|||||
QY      238 GCGATATATTAAAGAGATATACATCCATATAAATTTAAAGTGACGCTTTATTATAAA 297
      |||:|||||
Db      128 AlaValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 147
      |||:|||||
QY      298 AATATCATTCAGACGACGATGACGGGACGACATATAGACAGATCATCATCATAT 357
      :||:|||||

```

148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 167
 358 ACAGATAGACGCCGCTTCCCATTTGAAGAGATCAAGGATCAATCGACCGCAAGCAAGA 417
 168 GluAspArgAlaProValProPheGluGluValMetAspLysIleAsnAlaLysGlyVal 187
 418 TGTCTATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGACAGGGAT 477
 188 CysArgSerThrAlaLysTyrValArgAsnMetGluSerThrAlaPheHisArgAsp 207
 478 GCGGGAGAAACAAAGTCTTCTAAACCATCAAAATCAACACGCCCGAATCTAGGGCA 537
 208 AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly 227
 538 TGGCACACACTAATGAGACGTATACCGTGTGGGATCACCATGATATATCGAACGGCA 597
 228 TrpHisThrThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly 247
 598 ACCTCCGTCATTTGTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGTATTTCATAT 657
 248 ThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspGlu 267
 658 TTTGCAATGGCC 669
 268 PheValLeuAla 271

RESULT 14
 US-09-301-390-23
 ; Sequence 23, Application US/09301390
 ; GENERAL INFORMATION:
 ; APPLICANT: Rose, Timothy M.
 ; APPLICANT: Bosch, Marnix L.
 ; APPLICANT: Strand, Kurt
 ; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
 ; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morrison & Foerster
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/301,390
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/720,229
 ; FILING DATE: 26-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schiff, J. Michael
 ; REGISTRATION NUMBER: 40,253
 ; REFERENCE/DOCKET NUMBER: 29938-20002.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 885 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-301-390-23

Alignment Scores:
 US-09-301-390-23

Pred. No.: 2,32e-46 Length: 885
 Score: 504.00 Matches: 94
 Percent Similarity: 62.75% Conservative: 34
 Best Local Similarity: 46.08% Mismatches: 76
 Query Match: 42.00% Indels: 0
 DB: 17 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-09-301-390-23 (1-885)

QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGGTCGAGTGG 117
 DB 68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleLysAla 87
 QY 118 TCTGAGGAGAGTCTACGTTTATCTTGTCTCCCAACAGTGGTTCACCGTATCCGT 177
 DB 88 ArgAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValValGln 107
 QY 178 CTAGAACCCCGCGGAAATGTCGCACTAGAACCCAGTGGGTGAAGGAATC 237
 DB 108 PheGluGlnProArgProCysProArgAlaProAspGlyGlnAsnTyrThrGluGlyIle 127
 QY 238 GCGATATTATTAAGAGATATCATCTCCCAACAGTGGTTCACCGTATCCGT 177
 DB 128 AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 147
 QY 298 AATATCATTCAGACGACATGTCGAGGGGAGCAGATATAGACATCACTAATCGATAT 357
 DB 148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 167
 QY 358 ACAGATAGACGCGCGTTCATTCAGAGATACCGGATCAATCGACGGCAAGGA 417
 DB 168 GluAspArgAlaProValPheGluGluValMetAspLysIleAsnAlaLysGlyVal 187
 QY 418 TGCTCATCTAAAGCAAGTACCTTAGAACATGTATGTGAAGCGTTCACAGGAT 477
 DB 188 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHisArgAsp 207
 QY 478 GCGGAGAAAACAAGTACTCTTAAACCATCAAAATTCACACCGCGGATCAGGGA 537
 DB 208 AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly 227
 QY 538 TGGCAGACGACTAATGAGACGATACCGTGTGGGATCACCATGATATATCGACGGGA 597
 DB 228 TrpHisThrThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly 247
 QY 598 ACCTCGTCAATGTGTATAGAGGAATGGATGCCCGCTCTGTCTCCGTATTCATCAT 657
 DB 248 ThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspGlu 267
 QY 658 TTTGCAATGGCC 669
 DB 268 PheValLeuAla 271

RESULT 15

US-09-338-326-23
 ; Sequence 23, Application US/09338326
 ; GENERAL INFORMATION:
 ; APPLICANT: Rose, Timothy M.
 ; APPLICANT: Bosch, Marnix L.
 ; APPLICANT: Strand, Kurt
 ; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RBHV/KSHV
 ; NUMBER OF INVENTION: SUBFAMILY OF HERPES VIRUSES
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morrison & Foerster
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM: Floppy disk
 ; MEDIUM TYPE: IBM PC compatible
 ; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/338,326
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/720,229
 ; FILING DATE: 26-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schiff, J. Michael
 ; REGISTRATION NUMBER: 40,253
 ; REFERENCE/DOCKET NUMBER: 29938-20002.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 885 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-338-326-23

Alignment Scores:

Pred. No.: 2,32e-46 Length: 885
 Score: 504.00 Matches: 94
 Percent Similarity: 62.75% Conservative: 34
 Best Local Similarity: 46.08% Mismatches: 76
 Query Match: 42.00% Indels: 0
 DB: 17 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-09-338-326-23 (1-885)

QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGGTCGAGTGG 117
 DB 68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleLysAla 87
 QY 118 TCTGAGGAGAGTCTACGTTTATCTTGTCTCCCAACAGTGGTTCACCGTATCCGT 177
 DB 88 ArgAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValValGln 107
 QY 178 CTAGAACCCCGCGGAAATGTCGCACTAGAACCCAGTGGGTGAAGGAATC 237
 DB 108 PheGluGlnProArgProCysProArgAlaProAspGlyGlnAsnTyrThrGluGlyIle 127
 QY 238 GCGATATTATTAAGAGATATCATCTCCCAACAGTGGTTCACCGTATCCGT 297
 DB 128 AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 147
 QY 298 AATATCATTCAGACGACATGTCGAGGGGAGCAGATATAGACATCACTAATCGATAT 357
 DB 148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 167
 QY 358 ACAGATAGACGCGCGTTCATTCAGAGATACCGGATCAATCGACGGCAAGGA 417
 DB 168 GluAspArgAlaProValPheGluGluValMetAspLysIleAsnAlaLysGlyVal 187
 QY 418 TGCTCATCTAAAGCAAGTACCTTAGAACATGTATGTGAAGCGTTCACAGGAT 477
 DB 188 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHisArgAsp 207
 QY 478 GCGGAGAAAACAAGTACTCTTAAACCATCAAAATTCACACCGCGGATCAGGGA 537
 DB 208 AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly 227
 QY 538 TGGCAGACGACTAATGAGACGATACCGTGTGGGATCACCATGATATATCGACGGGA 597
 DB 228 TrpHisThrThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly 247
 QY 598 ACCTCGTCAATGTGTATAGAGGAATGGATGCCCGCTCTGTCTCCGTATTCATCAT 657

Db 248 ThrThrValAsnCysIleValGluValGluAlaArgSerValTyrProTyrAspGlu 267

Qy 658 TTTGCAATGGCC 669

Db 268 PheValLeuAla 271

Search completed: October 8, 2003, 19:15:29
Job time : 211.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 18:12:37 ; Search time 36.8116 Seconds
(without alignments)
5882.831 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200

Sequence: 1 atgcactatttttagcgga.....catattttgcattgccaat 672

Scoring table: BLOSUM62

Ygapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Ygapext 7.0
Delop 6.0 , Delext 7.0

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 1201306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154344_29960/app_query.fasta_1.5980
-DB=Published Applications_AA -OFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLAN=200 -THR_SCORE=ptc -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USPR=US09147052 @CGN_1_1_355 @runat_08102003_154344_29960
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

1	1200	100.0	1086	9	US-09-147-052-4	Sequence 4, Appli
2	498	41.5	904	9	US-09-894-998-18	Sequence 18, Appl
3	498	41.5	904	12	US-10-200-562-18	Sequence 18, Appl
4	498	41.5	904	12	US-10-237-551-18	Sequence 18, Appl
5	498	41.5	904	15	US-10-121-988-18	Sequence 18, Appl
6	314	26.2	62	15	US-10-131-591A-5	Sequence 5, Appli
7	314	26.2	456	9	US-09-147-052-2	Sequence 2, Appli
8	314	26.2	456	12	US-09-901-572A-3	Sequence 3, Appli
9	302	25.2	62	15	US-10-131-591A-6	Sequence 6, Appli
10	225	18.8	823	14	US-10-055-364-37	Sequence 37, Appl
11	219	18.2	844	14	US-10-055-364-40	Sequence 40, Appl
12	217.5	18.1	824	14	US-10-055-364-43	Sequence 43, Appl
13	217	18.1	818	14	US-10-055-364-44	Sequence 44, Appl
14	216.5	18.0	831	14	US-10-055-364-39	Sequence 39, Appl
15	213.5	17.8	808	14	US-10-055-364-38	Sequence 38, Appl
16	211.5	17.6	792	14	US-10-055-364-42	Sequence 42, Appl
17	205	17.1	865	14	US-10-055-364-24	Sequence 24, Appl
18	204	17.0	830	14	US-10-055-364-45	Sequence 45, Appl
19	202.5	16.9	907	15	US-10-223-538-2	Sequence 2, Appli
20	201.5	16.8	829	14	US-10-055-364-46	Sequence 46, Appl
21	196.5	16.4	791	14	US-10-055-364-41	Sequence 41, Appl
22	196.5	16.4	793	14	US-10-055-364-55	Sequence 55, Appl
23	85	7.1	405	10	US-09-919-497-63	Sequence 63, Appl
24	82	6.8	339	9	US-09-925-297-704	Sequence 704, Appl
25	79.5	6.6	316	15	US-10-042-894A-25	Sequence 25, Appl
26	79.5	6.6	1150	15	US-10-099-285-74	Sequence 74, Appl
27	79	6.6	441	10	US-09-866-572A-54	Sequence 54, Appl
28	79	6.6	441	10	US-09-866-570A-54	Sequence 54, Appl
29	79	6.6	441	15	US-10-166-984-54	Sequence 54, Appl
30	78.5	6.5	221	12	US-10-291-190-36	Sequence 36, Appl
31	78	6.5	2319	12	US-10-131-510A-6	Sequence 6, Appli
32	78	6.5	2319	15	US-10-187-319-6	GENERAL INFORMA
33	77.5	6.5	1328	15	US-10-083-357-1321	Sequence 1321, Ap
34	77	6.4	394	9	US-09-836-077-4	Sequence 4, Appli
35	77	6.4	694	10	US-09-935-868-18	Sequence 18, Appl
36	77	6.4	694	10	US-09-935-868-22	Sequence 22, Appl
37	77	6.4	694	12	US-10-282-162-18	Sequence 18, Appl
38	77	6.4	694	12	US-10-282-162-22	Sequence 22, Appl
39	77	6.4	694	15	US-10-287-035-18	Sequence 18, Appl
40	77	6.4	694	15	US-10-287-035-22	Sequence 22, Appl
41	76.5	6.4	430	11	US-09-443-745-23	Sequence 23, Appl
42	76.5	6.4	450	11	US-09-443-745-24	Sequence 24, Appl
43	76.5	6.4	548	12	US-09-846-589A-12	Sequence 12, Appl
44	76.5	6.4	691	10	US-09-935-868-20	Sequence 20, Appl
45	76.5	6.4	691	12	US-10-282-162-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

١٢

; AFFILIANT: FAISON, JOSEPH M.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

```

; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121.988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

Alignment Scores:
Pred. No.: 2,21e-46 Length: 904
Score: 498.00 Matches: 92
Percent Similarity: 63.73% Conservative: 38
Best Local Similarity: 45.10% Mismatches: 74
Query Match: 41.50% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-10-121-988-18 (1-904)
QY 58 ACTCATCTCCGAGTACCCAAATGTGACATCAAGAGACTTCTTTCGAGCGTCAGTGG 117
DB 82 ASNAIthrValAlaAlaaglyHisAlaThrLeuArgAlaHisLeuargGluIleLysVal 101
QY 118 TCTGAGGAAGAGTCTACGCTTTTATCTTTGTGCCACCAGTGGTTCACACCGTGATCCGT 177
DB 102 GluAsnAlaaspAlaGlnPheTyrValCysProProThrGlyAlaThrValGln 121
QY 178 CTAGAACCCCGCGAAATGTCCCGAACCTAGAAAGCCACCAGAGTGGGGTGAAGGAATC 237
DB 122 PheGluInProArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyIle 141
QY 238 GCGATATTATTAAGAGANATATCAGTCCATATAATTTAAAGTGACGCTTTATTATAAA 297
DB 142 AlavalValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 161
QY 298 AATATCATTTACAGACGACATATGACGGGGACGACATATAGACAGATCACTAATCGATAT 357
DB 162 AspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 181
QY 358 ACAGATAGGACCGCGTTTCCATTTCAAGAGATCAGGATCTATTCGACGCGCAAGGAAGA 417
DB 182 GluAspArgAlaProValProPheGluGluValIleaspLysIleasnThrLysGlyVal 201
QY 418 TGCTCATCTAAAGCAAGATACCTTAGAACCAATGTATGTGTGAAGCGTTTGACAGGGAT 477
DB 202 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgasp 221
QY 478 GCGGGAGAAAACAGACTTCTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGGCA 537
DB 222 AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArggly 241
QY 538 TGCACACGACACTAATGAGACGTATACCGTGTGGGGATCCACCATGATATATCAAGCGGA 597
DB 242 TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly 261
QY 598 ACCTCCGTCAAATGTATAGTAGGAATGGATGCCGCTCTGTGTTTCGTTATTCATAT 657
DB 262 ThrThrValAsnLysValGluGluValAspAlaArgSerValTyrProTyrAspGlu 281
QY 658 TTTTCAATGGCC 669
DB 282 PheValIleuAla 285

RESULT 6

```


Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.17% Indels: 0
DB: 9 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-09-147-052-2 (1-456)

QY 1 ATGCACATTTTAGCGGGAATTCATATTTTCCCTTAGTTATCTATATGTTAGTACGAAC 60
|||||
Db 1 MethisTyPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20
|||||
QY 61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
|||||
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
|||||
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACCAAGTGGTTCACACCGTGATCCGCTCTA 180
|||||
Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
|||||
QY 181 GAA 183
|||
Db 61 Glu 61

RESULT 8

US-09-901-572A-3

; Sequence 3, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

Alignment Scores:
Pred. No.: 5.57e-26 Length: 456
Score: 314.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.17% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-09-901-572A-3 (1-456)

QY 1 ATGCACATTTTAGCGGGAATTCATATTTTCCCTTAGTTATCTATATGTTAGTACGAAC 60
|||||
Db 1 MethisTyPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20
|||||
QY 61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
|||||
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
|||||
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACCAAGTGGTTCACACCGTGATCCGCTCTA 180
|||||
Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
|||||
QY 181 GAA 183
|||
Db 61 Glu 61

RESULT 9

US-10-131-591A-6

; Sequence 6, Application US/10131591A
; Publication No. US20030059799A1

; GENERAL INFORMATION:

; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified VgB signal
US-10-131-591A-6

Alignment Scores:

Pred. No.: 6.34e-25 Length: 62
Score: 302.00 Matches: 59
Percent Similarity: 96.72% Conservative: 0
Best Local Similarity: 96.72% Mismatches: 2
Query Match: 25.17% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x US-10-131-591A-6 (1-62)

QY 1 ATGCACATTTTAGCGGGAATTCATATTTTCCCTTAGTTATCTATATGTTAGTACGAAC 60
|||||
Db 1 MethisTyPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrGln 20
|||||
QY 61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
|||||
Db 21 SerSerProSerThrGlnGlnValThrSerArgGluValValSerSerValGlnLeuSer 40
|||||
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACCAAGTGGTTCACACCGTGATCCGCTCTA 180
|||||
Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
|||||
QY 181 GAA 183
|||
Db 61 Glu 61

RESULT 10

US-10-055-364-37

; Sequence 37, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Human herpesvirus 8
US-10-055-364-37

Alignment Scores:

Pred. No.: 5.55e-16 Length: 823
Score: 225.00 Matches: 67
Percent Similarity: 46.95% Conservative: 33
Best Local Similarity: 31.46% Mismatches: 99
Query Match: 18.75% Indels: 14
DB: 14 Gaps: 5

```
US-09-147-052-3_COPY_1_672 (1-672) x US-10-055-364-37 (1-823)
QY 49 TATGTAGCAACTCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGC 108
   :::::|||||:::
Db 31 PheGlnThrSerSerSerProThrProGlySerSerLys-----AlaPro 47
QY 109 GTCCAGTTCTCTGAGGAGAGTCT-----ACGTTTATCTT 144
   |||||:::
Db 48 ThrLysProGlyGluGluAlaSerGlyProLysSerValAspPheTyrGlnPheArgVal 67
QY 145 TGTCCCCACCAGTGGTTCACCGTGTCTAGAACCGCCGAAATGTCCCGAA 204
   :::::|||||:::
Db 68 CysSerAlaSerIleThrGlyLeuPheArgPheAsnLeuGluGlnThrCysProAsp 87
QY 205 CCTAGAAAGCCAGGTGGGTGAAGATCGCGATATTATTTAAAGAGAATATCAGT 264
   |||||:::
Db 88 ---ThrLysAspLysTyrHisGlnGluGlyIleLeuLeuValTyrLysAsnIleVal 106
QY 265 CCATATAAATTTAAAGTGCAGCGTTTATTATATAAAATATCATTCAGACGACGACG 324
   |||||:::
Db 107 ProHisIlePheLysValArgArgTyrArgLysIleAlaThrSerValThrValTyrArg 126
QY 325 GGCAGACATATAGACAGACTCAATCATATACAGATAGGACCGCGTTTCCATTGAA 384
   |||||:::
Db 127 GlyLeuThrGluSerAlaIleThrAsnLysTyrGluLeuProArgProValProLeuTyr 146
QY 385 GAGATCAGGATCTAATCGAGGCAAGAGATGCTCATCTAAGCAAGATACCTTAGA 444
   |||||:::
Db 147 GluIleSerHisMet---AspSerThrTyrGlnCysPheSerSerMetLysValAsnVal 165
QY 445 AACATGTATATGTTGAAGCGTTTCACAGGGATCGGGAGAAAACAAGTACTCTTCTAAA 504
   |||||:::
Db 166 AsnGlyValGluAsnThrPheThrAspArgAspValAsnThrThrValPheLeuGln 185
QY 505 CCATCAAAATTAACACCCCGCAATCTAGGCGATGGCACACGACTAATGAGACGATACC 564
   |||||:::
Db 186 ProValGluGlyLeuThrAspAsnIleGlnArgTyrPheSerGlnProValIleTyrAla 205
QY 565 GNGTGGGA---TCACCATGATATATCGAACGGGAACCTCCGTCAAATGTATAGTAG 621
   |||||:::
Db 206 GluProGlyTyrPheProGlyIleTyrArgValArgThrThrValAsnCysGluIleVal 225
QY 622 GAAATGGATGCCCGCTGCTGTTCCGTATTCATATTTT 660
   :::::|||||:::
Db 226 AspMetIleAlaArgSerAlaGluProTyrAsnTyrPhe 238
```

RESULT 11

```
US-10-055-364-40
; Sequence 40, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
```

```
; FILE REFERENCE: 61750-379
```

```
; CURRENT APPLICATION NUMBER: US/10/055,364
```

```
; CURRENT FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: US/09/612,204
```

```
; PRIOR FILING DATE: 2000-07-07
```

```
; PRIOR APPLICATION NUMBER: US/60/142,736
```

```
; PRIOR FILING DATE: 1999-07-08
```

```
; PRIOR APPLICATION NUMBER: US/60/168,532
```

```
; PRIOR FILING DATE: 1999-12-02
```

```
; NUMBER OF SEQ ID NOS: 55
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 40
```

```
; LENGTH: 844
```

```
; TYPE: PRT
```

```
; ORGANISM: Bovine herpesvirus 4
```

```
US-10-055-364-40
```

```
Alignment Scores:
```

```
Pred. No.: 2.61e-15 Length: 844
```

```
Score: 219.00 Matches: 69
Percent Similarity: 45.97% Conservative: 28
Best Local Similarity: 32.70% Mismatches: 106
Query Match: 18.25% Indels: 8
DB: 14 Gaps: 6
```

```
US-09-147-052-3_COPY_1_672 (1-672) x US-10-055-364-40 (1-844)
```

```
QY 55 ACGAACCTCATCTCCGAGTACCCAAAATGTG---ACATCAAGAGAAGTTGTTTCGAGCGTC 111
   :::::|||||:::
Db 62 SerThrSerLysProSerThrAspAsnGlnGlyThrSerThrProThrIleProThrVal 81
QY 112 CAGTTGCTGAGGAGAGACTCTACGTTTAT-----CTTTGTCGCCACCAAGTGGT 162
   |||||:::
Db 82 ThrAspAspThrAlaSerLysAsnPheTyrLysTyrArgValCysSerAlaSerSerSer 101
QY 163 TCA---ACCGTGATCCGCTAGAACCGCGGAAATGTCCGAACTAGAAAAGCCACC 219
   |||||:::
Db 102 SerGlyGluLeuPheArgPheAspLeuAspGlnThrCysProAsp---ThrLysAspLys 120
QY 220 GAGTGGGTGAAGGAATCGCGATATTATTTAAAGAGAATATCATGCTCCATATAAATTTAAA 279
   :::::|||||:::
Db 121 LysHisValGluGlyIleLeuLeuValLeuLysLysAsnIleValProTyrIlePheLys 140
QY 280 GTGACGCTTTATTATAAAATATCATTCAGACGACGACATGCGCGGGGACACATATAGA 339
   |||||:::
Db 141 ValArgLysTyrArgLysIleAlaThrSerValThrValTyrArgGlyTyrPheSerGlnAla 160
QY 340 CAGATCACTAATCATATACAGATAGGACCGCGGTTTCCATTGAAGAGATCAGCGATCTA 399
   :::::|||||:::
Db 161 AlaValThrAsnArgAspAspIleSerArgAlaIleProTyrAsnGluIleSer---Met 179
QY 400 ATCGACGCGCAAGGAAGATGCTCATCTAAACAGAGATACCTTGAACAACATCAAAATTCAC 459
   |||||:::
Db 180 IleAspArgThrTyrHisCysPheSerAlaMetAlaThrValIleAsnGlyIleLeuAsn 199
QY 460 GAAGCGTTTGACAGGATCGCGGAGAAAACAAGTACTCTTAAACCATCAAAATTCAC 519
   |||||:::
Db 200 ThrTyrIleAspArgAspSerGluAsnLysSerValProLeuGlnProValAlaGlyLeu 219
QY 520 ACGCCCGAATCTAGGCGATGGCACACGACTAATGAGACGATACCGTGTGGGGA---TCA 576
   |||||:::
Db 220 ThrGluAsnIleAsnArgTyrPheSerGlnProLeuIleTyrAlaGluProGlyTyrPhe 239
QY 577 CCATGGATATATCGAACGGGAACCTCCGTCAAATGTATAGTAGAGAAATGATGCCCGC 636
   |||||:::
Db 240 ProGlyIleTyrArgValArgThrThrValAsnCysGluValValAspMetTyrAlaArg 259
QY 637 TCTGTGTTTCGCTATTCATATTTTCAATGGCC 669
   |||||:::
Db 260 SerValGluProTyrThrHisPheIleThrAla 270
```

RESULT 12

```
US-10-055-364-43
```

```
; Sequence 43, Application US/10055364
```

```
; Publication No. US20020155433A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Patience, Clive
```

```
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
```

```
; FILE REFERENCE: 61750-379
```

```
; CURRENT APPLICATION NUMBER: US/10/055,364
```

```
; CURRENT FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: US/09/612,204
```

```
; PRIOR FILING DATE: 2000-07-07
```

```
; PRIOR APPLICATION NUMBER: US/60/142,736
```

```
; PRIOR FILING DATE: 1999-07-08
```

```
; PRIOR APPLICATION NUMBER: US/60/168,532
```

```
; PRIOR FILING DATE: 1999-12-02
```

```
; NUMBER OF SEQ ID NOS: 55
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 43
```

```
; LENGTH: 824
```

```
; TYPE: PRT
```

; ORGANISM: Equine herpesvirus 2
US-10-055-364-43

Alignment Scores:

Pred. No.: 3,8e-15 Length: 824
Score: 217,50 Matches: 63
Percent Similarity: 46,50% Conservative: 30
Best Local Similarity: 31,50% Mismatches: 102
Query Match: 18,12% Indels: 5
DB: 14 Gaps: 5

US-09-147-052-3_COPY_1_672 (1-672) x US-10-055-364-43 (1-824)

```
Qy 70 AGTACCAAAATGTGACATCAAGA---GAAGTGTGTTTCGAGCGTCCAGTGTCTGTCAGGAA 126
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 ThrThrProPheAlaThrHisArgProGluValValAlaGluGluAsnProAlaAsnPro 52
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 127 GAGTCTACGTTTATCTTTGT---CCCCACCGAGTGGTCAACCGTGATCCGCTAGAA 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 PheLeuProPheArgValCysGlyAlaSerProThrGlyGlyGluIlePheArgPhePro 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184 CCGCGCGAAATGTCGCCGACCTAGAAAGCCAGGAGTGGGTGAAGGAATCGCGATA 243
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 LeuGluGluSerCysProAsn---ThrGluAspLysAspHisIleGluGlyIleAlaLeu 91
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244 TTATTTAAAGAGATATCAGTCCATATAAATTTAAAGTACGCGTCTTATTTATAAATATC 303
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 IleTyrLysThrAsnIleValProTyrValPheAsnValArgLysTyrArgLysIleMet 111
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 ATTCAGACGACGACGACGCGGACGACATATAGACAGATCACTAATCGATATACAGAT 363
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 ThrSerThrThrIleTyrLysGlyTrpSerGluAspAlaIleThrAsnGlnHisThrArg 131
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 364 AGGACGCGCGTTTCATTTGAAGAGATCACGGATCAATACGACGCAAGGAAGATGCTCA 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 SerTyrAlaValProLeuTyrGluVal---GlnMetMetAspHisTyrTyrGlnCysPhe 150
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 424 TCTAAACGACAGATACCTTAGAACAATGTATATGTTCAAGCGTTTTCAGAGGATCGCGGA 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 SerAlaValGlnValAsnGluGlyGlyHisValAsnThrTyrTyrAspArgAspGlyTyr 170
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 484 GAAAAACAAGTACTCTTAAACCATCAAAATTCACACGCGCGCAATCTAGGCGATGCGAC 543
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 AsnGlnThrAlaPheLeuLysProAlaAspGlyLeuThrSerSerIleThrArgTyrGln 190
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 544 ACAGCTAATGAGAGGTAT---ACCGTGTGGGGATCACCATGGATATATCAACGCGGAACC 600
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 SerGlnProGluValTyrAlaThrProArgAsnLeuLeuTrpSerTyrThrThrArgThr 210
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 601 TCCGTCATTTCTATAGTAGAGGAATGGATGCGCGCTGTGTTCGTTTCGTTATTCATATTTT 660
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 ThrValAsnCysGluValThrGluMetSerAlaArgSerMetLysProPheGluPhePhe 230
```

RESULT 13

US-10-055-364-44
; Sequence 44, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/169,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818

; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
US-10-055-364-44

Alignment Scores:

Pred. No.: 4,31e-15 Length: 818
Score: 217,00 Matches: 61
Percent Similarity: 45,26% Conservative: 25
Best Local Similarity: 32,11% Mismatches: 96
Query Match: 18,08% Indels: 8
DB: 14 Gaps: 4

US-09-147-052-3_COPY_1_672 (1-672) x US-10-055-364-44 (1-818)

```
Qy 121 GAGGAAGAGCTACG-----TTTTTCTTTGTCCTCCACACAGTGGGTCA 165
   ||||| :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46 GluAsnThrThrAsnProPheLeuLeuPheArgValCysGlyAlaSerProThrGly 65
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 166 ACCGTGATCCGTCTAGAACCCCGCGAAATATGCCGAACTAGAAAAGCCACCGAGTGG 225
   :::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 GluIlePheArgPheProLeuGluGluAsnCysProAsn---ThrGluAspLysGluHis 84
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 226 GGTCAAGGAATCGGATATATTATAAGAGAAATATCAGTCCATATAAATTTAAAGTGACG 285
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 ValGluGlyIleLeuLeuIleTyrLysThrAsnIleValProTyrIlePheAsnValArg 104
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 286 CTTTATTATAAATATCATTCAGACGACGACATGCGGGGAGCAGCATATAGACAGATC 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 LysTyrArgLysLeuValThrSerThrThrIleTyrLysGlyTrpSerGlnAspAlaIle 124
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 346 ACTAATCGATATACAGATAGACCCCGTTTCCATTTGAAGAGATCAGGATCTAATCGAC 405
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 ThrAsnGlnTyrThrSerSerPheAlaMetProLeuTrpGlu---AlaArgLeuValAsp 143
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 406 GGCAAAAGGAAGATCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATTTGAACGC 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 TyrAsnTyrGluCysTyrAsnGlyIleGlnValThrGluAsnGlyHisLeuThrTyr 163
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 466 TTTGACAGGATCGGGAGAAAACAAGTACTTCTTAAACCATCAAAATTCACACGCGCC 525
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 ValAspArgAspGlyTyrAsnGluSerValArgLeuValProAlaAspGlyLeuThrSer 183
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 526 GAATCTAGGCGATGCGACACGACGATGACGATAT---ACCGTGTGGGGATCACCATGG 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 SerIleArgArgTyrHisSerGlnProGluLeuTyrValThrProArgAsnLeuLeuTrp 203
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 583 ATATATCGAAGCGGAACCTCCGTCATTTGTATAGAGGAAATGGATGCCGCTCTGTG 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 SerTyrThrThrArgThrThrValAsnCysGluValIleAspMetThrAlaArgSerHis 223
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 643 TTTCCGTTATTCATATTTTGCATATGCGCAAT 672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 LysProPheGluTyrPheValThrAlaSer 233
   :::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 14

US-10-055-364-39
; Sequence 39, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/169,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39

```

; LENGTH: 831
; TYPE: PRT
; ORGANISM: Murine herpesvirus 68
; US-10-055-364-39

Alignment Scores:
Pred. No.:      4,93e-15      Length:      831
Score:          216.50      Matches:      58
Percent Similarity: 47.49%      Conservative: 27
Best Local Similarity: 32.40%      Mismatches: 91
Query Match:      18.04%      Indels:      3
DB:              14      Gaps:      3

US-09-147-052-3_COPY_1_672 (1-672) x US-10-055-364-39 (1-831)
QY 136 TTTTATCTTTTCCCCACCAGTGGGTCAACGCTGATCGTCTAGAACGCCGCGGAAAA 195
Db 67 PheArgValCysGlyValAlaAlaThrGlyCluThrPheArgPheAspLeuAspLysThr 86
QY 196 TGTCCCGAACCTAGAAAAGCCACCGAGTGGGGTGAAGGAATCGGATATATTATAAAGAG 255
Db 87 CysProSerThrGlnAsp---LysLysHisValGluGlyIleLeuLeuValTyrLysIle 105
QY 256 ATATATCAGTCCATATAAATTTAAAGTACGCTTTATTTATAAAATATCATTCAGACGAG 315
Db 106 AsnIleValProTyrIlePheLysIleGArgGlyrArgLysIleIleThrGlnLeuThr 125
QY 316 ACATGGAGGGGACGACATATAGACATCACTAATCGATATACATAGGACGCCCGTT 375
Db 126 IleTPrArgGlyLeuThrThrSerSerValThrGlyLysPheGluMetAlaThrGlnAla 145
QY 376 TCATTGAGAGATCAGCGATCTAATCGACGGCAAGGAAGATGCTCATCTAAAGCAAGA 435
Db 146 HisGluTPrGluValGlyAsp---PheAspSerIleTyrGlnCysTyrAsnSerAlaThr 164
QY 436 TACCTTAGAACAACTGATATGTTGAACGCTTGACAGGGATCGGGGAGAAAACAAGTA 495
Db 165 MetValValAsnAsnValArgGlnValTyrValAspArgAspGlyValAsnLysThrVal 184
QY 496 CTTCTAAACACTCAAAAATTCACACGCCCGCAATCTAGGCGCATCGCACGACTAATGAG 555
Db 185 AsnIleArgProValAspGlyLeuThrGlyAsnIleGlnArgTyrPheSerGlnProThr 204
QY 556 AGCTATACCGTGTGGGGA---TCACCATGGATATATCGAACGGGAACCTCCGTCATTTGT 612
Db 205 LeuTyrSerGluProGlyTyrMetProGlyPheTyrArgValArgThrThrValAsnCys 224
QY 613 ATAGTAGAGAAAATGGATGCCCGCTGTGTTCCGTATTTCATATTGTCATGGCC 669
Db 225 GluIleValAspMetValAlaArgSerMetAspProTyrAsnTyrIleAlaThrAla 243

RESULT 15
US-10-055-364-38
; Sequence 38, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: patentin version 3.0
; SEQ ID NO 38
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Rhesus monkey rhadinovirus

```

Search completed: October 8, 2003, 19:27:44
Job time : 43.8116 secs

10

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:41:47 ; Search time 10.5802 Seconds
(without alignments)
5374.731 Million cell updates/sec

Title: US-09-147-052-3_COPY_1_672

Perfect score: 1200

Sequence: 1 atgcactattttgagcgaa.....catattttgcaatggccaat 672

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xmlp

-MODEL=frame_n2p.model
-Q=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154342_29880/app_query.fasta_1.5980
-DB=Issued Patents_AA -OFFMT=fastan -SURFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_126@runat_08102003_154342_29880 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1200	100.0	865	1	US-07-803-633A-13
2	659.5	55.0	879	1	US-08-220-151-2
3	659.5	55.0	879	1	US-08-220-151-3
4	659.5	55.0	879	1	US-08-413-118-2
5	659.5	55.0	879	1	US-08-413-118-3
6	659.5	55.0	879	1	US-08-413-118-106
7	659.5	55.0	879	3	US-08-473-446-2
8	659.5	55.0	879	3	US-08-473-446-3
9	659.5	55.0	879	3	US-08-473-446-106
10	614.5	51.2	1041	1	US-08-220-151-4
11	614.5	51.2	1041	1	US-08-413-118-4
12	614.5	51.2	1041	3	US-08-473-446-4

13	609	50.7	913	3	US-09-232-468A-2	Sequence 2, Appl
14	609	50.7	913	4	US-09-784-984B-49	Sequence 49, Appl
15	606	50.5	868	1	US-08-220-151-7	Sequence 7, Appl
16	606	50.5	868	1	US-08-413-118-7	Sequence 7, Appl
17	606	50.5	868	3	US-08-804-439A-21	Sequence 21, Appl
18	606	50.5	868	3	US-08-473-446-7	Sequence 7, Appl
19	606	50.5	868	3	US-08-720-229-21	Sequence 21, Appl
20	605	50.4	913	1	US-08-220-151-6	Sequence 6, Appl
21	605	50.4	913	1	US-08-413-118-6	Sequence 6, Appl
22	605	50.4	913	3	US-08-473-446-6	Sequence 6, Appl
23	605	50.4	913	6	5196516-8	Patent No. 5196516
24	583	48.6	943	3	US-08-911-321-4	Sequence 4, Appl
25	582	48.5	933	2	US-08-682-847-2	Sequence 2, Appl
26	525.5	43.8	980	1	US-08-220-151-5	Sequence 5, Appl
27	525.5	43.8	980	1	US-08-413-118-5	Sequence 5, Appl
28	525.5	43.8	980	3	US-08-473-446-5	Sequence 5, Appl
29	513	42.8	891	1	US-08-042-747A-6	Sequence 6, Appl
30	504	42.0	885	1	US-08-042-747A-8	Sequence 8, Appl
31	504	42.0	885	3	US-08-804-439A-23	Sequence 23, Appl
32	504	42.0	885	3	US-08-720-229-23	Sequence 23, Appl
33	504	42.0	903	1	US-08-220-151-8	Sequence 8, Appl
34	504	42.0	903	1	US-08-413-118-8	Sequence 8, Appl
35	504	42.0	903	3	US-08-804-439A-22	Sequence 22, Appl
36	504	42.0	903	3	US-08-473-446-8	Sequence 8, Appl
37	504	42.0	903	3	US-08-720-229-22	Sequence 22, Appl
38	504	42.0	904	3	US-08-632-537-1	Sequence 1, Appl
39	504	42.0	904	5	PCT-US96-05316-1	Sequence 1, Appl
40	504	42.0	904	6	5244792-4	Patent No. 5244792
41	500	41.7	904	3	US-08-632-537-2	Sequence 2, Appl
42	500	41.7	904	5	PCT-US96-05316-2	Sequence 2, Appl
43	500	41.7	904	6	5244792-3	Patent No. 5244792
44	498	41.5	904	4	US-09-894-998A-18	Sequence 18, Appl
45	225	18.8	845	3	US-08-804-439A-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1
US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025ofu
; APPLICANT: OGAWA, Ryohel
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848

128	Db	128	TyAlaValIleThrAsnArgTyThrAspArgValProIleGlyValProGluIleThr	147
394	Qy	394	GATCTAATCAGCGCAAGAAAGTGTCTATCAAGCAAGCATACCTTAGAACCAATGTA	453
148	Db	148	GluLeuIleAspArgArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr	167
454	Qy	454	TATGTTGAAGCGTTTGACGGGATGCGGGAGAAACAAAGTACTTCTAAAACCATCAAAA	513
168	Db	168	GlupheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys	187
514	Qy	514	TTCAACACGCCCGAATCTAGGGCATGGCACACGACTAATGACGCTATACCGTGGGGGA	573
188	Db	188	PheAsnThrProGlySerArgGlyTyrPheHisThrValAsnAspThrTyrThrLysIleGly	207
574	Qy	574	TCACCATGGATATCAAGACGGGAACCTCCGCTCAATTGTATAGTAGACGAAATGATGCC	633
208	Db	208	GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla	227
634	Qy	634	CGCTCTCTGTTTCGGTATTCATATTTTTCGAATGCC	669
228	Db	228	ArgSerValTyrProTyrAspSerPheAlaIleSer	239

```

RESULT 3
US-08-220-151-3
: Sequence 3, Application US/08220151
: Patent No. 5529780
: GENERAL INFORMATION:
: APPLICANT: Paolletti, Enzo
: APPLICANT: Limbach, Keith J.
: TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
: TITLE OF INVENTION: CANINE HERPESVIRUS qB, qC AND qD AND USES THEREFOR

```

NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtiss, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036

```

?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
?
? COMPUTER: IBM PC compatible
?
? OPERATING SYSTEM: PC-DOS/MS-DOS
?
? SOFTWARE: Patent In Release #1.0, Version #1.25
?
? CURRENT APPLICATION DATA:
?
? APPLICATION NUMBER: US/08/220,151
?
? FILING DATE: 30-MAR-1994
?
?
```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712

```

; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 679 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-320-151-3

```

Alignment Scores:	
Pred. No.:	2,12e-70
Score:	659.50
Percent Similarity:	68.10%
Best Local Similarity:	53.88%
Query Match:	54.96%
Length:	879
Matches:	125
Conservative:	33
Mismatches:	57
Indels:	17

DB:	1	Gaps:	2
US-09-147-052-3_COPY_1_672 (1-672) x US-08-220-151-3 (1-879)			
QY	25 ATATTTTCTCTATAGTATATCTA-----		----- 48
Db	8 IlePhePheIleIleIleIleIleIleIleIleCysAspProThrThrProGluSerThrIle 27		
QY	49 -----TATGTCAGCAACTCATCTCCGAGTACCCAAAAATGTGACATCAAGAGAA 96		
Db	28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspAspIleArgGlu 47		
QY	97 GTGTGTTCCGAGCGGCCAGTGTG---TCGTAGGAGAGTCTACGTTTTTATCTTTGTGCCCA 153		
Db	48 IleLeuArgGluSerGluIleGluSerAspThrSerThrPheTyrMetCysProPro 67		
QY	154 CCAGTGGGTTCACACCGTGATCCGCTAGAACCGCGCGGAAATCTCCGAACTTAGAAA 213		
Db	68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu 87		
QY	214 GCCACCGAGTGGGGTGAAGGAATCGCATATATTAAAGAGAATATCAGTCCCATATAAA 273		
Db	88 GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys 107		
QY	274 TTTAAAGTGCAGCTTTATTATAAAATATCATTCAGACGACGATGGACGGGACGACA 333		
Db	108 PheLysAlaAsnIleTyrTyrLysAsnIleIleIleThrThrValTyrPserGlySerThr 127		
QY	334 TATAGACAGATCACTAATCCGATATACAGATGAGGACGCCGCTTCCATTAAGAGATCACG 393		
Db	128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147		
QY	394 GATCTAATCAGCGGCAAGGAAGTGTCTATCAAGCAAGATACCTTAGAAACAATGTA 453		
Db	148 GluLeuIleAspArgArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnThr 167		
QY	454 TATGTTGAAGGTTTGACAGGGATGCGGGAGAAAACAAGTACTCTTAAACCATCAAAA 513		
Db	168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187		
QY	514 TTCAACACGCGCGGAATCTAGGCGATGGCACACGACTAATGAGACGTATACCGTGGGGA 573		
Db	188 PheAsnThrProGlySerArgGlyTyrPheHisThrValAsnAspThrTyrThrLysIleGly 207		
QY	574 TCACCATGGATATCAACGCGGGAACCTCCGCTCAATTGTATAGTAGAGGAATGGATGCC 633		
Db	208 GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla 227		
QY	634 CGCTCTGTGTTTCGGTATTCATATTTTGCATGCCC 669		
Db	228 ArgSerValTyrProTyrAspSerPheAlaIleSer 239		

```

RESULT 4
US-08-413-118-2
: Sequence 2, Application US/08413118
: Patent No. 5688920
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: APPLICANT: LIMBACH, KEITH J.
: TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
: TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
: NUMBER OF SEQUENCES: 128
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
: STREET: 530 FIFTH AVENUE, 25TH FLOOR
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

!

88	GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys	107
274	TTTAAAGTCAGCGCTTTATTATAAAATATCATTCAGACGACGATCGAGCGGGGACGACA	333
108	PheLysAlaAsnIleTyrTyrLysAsnIleIleIleThrValTrpSerGlySerThr	127
334	TATACACAGATCATCATTCGATATACAGATAGGACGCCCGTTTCCATTGAAGAGATCAGC	393
128	TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr	147
394	GATCTAAATCGACGGCAGGAGAGATGCTCATCTAAACGACAGATACCTTAGAACATGTGA	453
148	GluLeuIleAspArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr	167
454	TATGTTGAAGCGTTTGACAGGGATGCGGGGAGAAACACAGTACTTCTAAACCATCAAAA	513
168	GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys	187
514	TTCAACACGCCCGGAATCTAGGCGATGCCACACCACTAATGACAGCTATACCGTGTGGGGA	573
198	PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly	207
574	TCACCATGGATATATCAAGACGGGAACCTCCGCTCAATTGTATAGTAGAGAAATGGATGCC	633
208	GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla	227
634	CGCTCTGTGTTCCGTATTCATATTMTTGAATGCC	669
228	ArgSerValTyrProTyrAspSerPheAlaIleSer	239

```

; FRAGMENT TYPE: N-terminal
US-08-473-446-3

Alignment Scores:
Pred. No.: 2,12e-70 Length: 879
Score: 659.50 Matches: 125
Percent Similarity: 68.10% Conservative: 33
Best Local Similarity: 53.88% Mismatches: 57
Query Match: 54.96% Indels: 17
DB: 3 Gaps: 2

US-09-147-052-3_COPY_1_672 (1-672) x US-08-473-446-3 (1-879)

QY 25 ATATTTTCCTTATAGTTATCTA----- 48
Db 8 IlePhePheIleIleTyrThrLeuIleIleCysAspProThrThrProGluSerThrIle 27
QY 49 -----TATGTCAGAACTCATCTCCGAGTACCCAAAATGTGACATCAAGAA 96
Db 28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspAspIleArgGlu 47
QY 97 GTTGTTCGAGCGCCAGTTG---TCTGAGAAGAGTCTAGCTTTTATCTTTGTCCCA 153
Db 48 IleLeuArgGluSerGlnIleGluSerAspThrSerThrPheTyrMetCysProPro 67
QY 154 CCAGTGGTTCACCGTATCCGTCCTAGAACCGCGGAAATGTCCGACCTAGAAA 213
Db 68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu 87
QY 214 GCCACCGAGTGGGTGAAGGAATCCGATATATTAAAGAGAATATCAGTCCATATA 273
Db 88 GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys 107
QY 274 TTTAAAGTGACGCTTTATTATAAAATATCATTACAGACGACCATGACGGGACGACA 333
Db 108 PheLysAlaAsnIleTyrTyrLysAsnIleIleIleThrValThrValTrpSerGlySerThr 127
QY 334 TATAGACAGTCACTAATCGATATACAGATAGACGCGCGTTCATTCATTCAGAGATCAG 393
Db 128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147
QY 394 GATCTAATCGAGCGGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTA 453
Db 148 GluLeuIleAspArgArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnThr 167
QY 454 TATGTTGAAGCGTTTGACAGGATCGCGGAGAAAACAAGTACTTAAACCACTCAAAA 513
Db 168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187
QY 514 TTCACACGCGCGGATCTAGGCGATGGCACGACTAATGACGATATACCGTGTGGGA 573
Db 188 PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrThrLysIleGly 207
QY 574 TCACCATGGATATCGAACGGGAACCTCCGTCATATTGTATAGTAGAGGAATGGATGCC 633
Db 208 GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla 227
QY 634 CGCTCTGTGTTCCGATATTCATATTTCGAATGCC 669
Db 228 ArgSerValTyrProTyrAspSerPheAlaIleSer 239

RESULT 9
US-08-473-446-106
; Sequence 106, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.

```

```

; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
; US-08-473-446-106

```

```

Alignment Scores:
Pred. No.: 2,12e-70 Length: 879
Score: 659.50 Matches: 125
Percent Similarity: 68.10% Conservative: 33
Best Local Similarity: 53.88% Mismatches: 57
Query Match: 54.96% Indels: 17
DB: 3 Gaps: 2

US-09-147-052-3_COPY_1_672 (1-672) x US-08-473-446-106 (1-879)

```

```

QY 25 ATATTTTCCTTATAGTTATCTA----- 48
Db 8 IlePhePheIleIleTyrThrLeuIleIleCysAspProThrThrProGluSerThrIle 27
QY 49 -----TATGTCAGAACTCATCTCCGAGTACCCAAAATGTGACATCAAGAA 96
Db 28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspAspIleArgGlu 47
QY 97 GTTGTTCGAGCGCCAGTTG---TCTGAGAAGAGTCTAGCTTTTATCTTTGTCCCA 153
Db 48 IleLeuArgGluSerGlnIleGluSerAspThrSerThrPheTyrMetCysProPro 67
QY 154 CCAGTGGTTCACCGTATCCGTCCTAGAACCGCGGAAATGTCCGACCTAGAAA 213
Db 68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu 87
QY 214 GCCACCGAGTGGGTGAAGGAATCCGATATATTAAAGAGAATATCAGTCCATATA 273
Db 88 GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys 107
QY 274 TTTAAAGTGACGCTTTATTATAAAATATCATTACAGACGACCATGACGGGACGACA 333
Db 108 PheLysAlaAsnIleTyrTyrLysAsnIleIleIleThrValThrValTrpSerGlySerThr 127
QY 334 TATAGACAGTCACTAATCGATATACAGATAGACGCGCGTTCATTCATTCAGAGATCAG 393
Db 128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147
QY 394 GATCTAATCGAGCGGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTA 453

```

```

Db 148 GluLeuIleAspArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnSerTyr 167
      ::::::::::::::::::::: ||| ||||||||| |||:::|||||
Qy 454 TATGTTGAACGGTTGACAGGGATGCGGGAGAAACAAAGTACTTCTAAACACCAAAA 513
      |||||||::: ||| |||||||
Db 168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187
      ::::::::::::::::::::: ||| |||||||
Qy 514 TTCAACACGCCGATCTAGCGATGCCACACCACTAATGAGAGGTATACCGTGGGGA 573
      |||||||::: ||| |||||||
Db 188 PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly 207
      |||||||::: ||| |||||||
Qy 574 TCACCATGGATATATCAACGGCACTCCGTCATTTGATAGTACAGGAAATGGATGCC 633
      ||| ::::::::::::::::::::: ||| |||||||
Db 208 GlySerGlyPheThrHisSerGlyThrSerValAsnCysIleValGluValAspAla 227
      ||| ::::::::::::::::::::: ||| |||||||
Qy 634 CGCTCTGTGTTCCGTATTCATATTTTGGCAATGGCC 669
      |||||||::: ||| |||||||
Db 228 ArgSerValTyrProTyrAspSerPheAlaIleSer 239
      |||||||::: ||| |||||||

```

RESULT 10

```

US-08-220-151-4
; Sequence 4, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTWS

```

INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-4

```

Alignment Scores:

```

Pred. No.: 5,97e-65 Length: 1041
Score: 614.50 Matches: 129
Percent Similarity: 44.92% Conservative: 39
Best Local Similarity: 34.49% Mismatches: 53
Query Match: 51.21% Indels: 153
DB: 1 Gaps: 4

```

US-09-147-052-3_COPY_1_672 (1-672) x US-08-220-151-4 (1-1041)

```

Qy 7 TATTATTAGCGGAATTCATATTT----- 30
      |||||||::: ||| |||
Db 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
      ::::::::::::::::::::: ||| |||
Qy 31 -----TTCTTTATAGTT 42
      ::::::::::::::::::::: |||
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
      ::::::::::::::::::::: |||
Qy 43 ATTCTATAT----- 51
      ::::::::::::::::::::: |||
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyClnSerGlySerThrSerGluGln 81
      ::::::::::::::::::::: |||
Qy 52 -----GGTACGAACCTCATCTCCGAGTACCCAA 78
      |||||
Db 82 ProArgArgThrValAlaThrProGluValGlyThrProProLysProThrThrAsp 101
      |||||
Qy 79 AATGTGACATCAAGAGAAGTTGTTGAGCGTCCAGTTGCTCGAGGAAG----- 129
      ::::::: |||
Db 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
      ::::::: |||
Qy 130 ---TCTACGTTTATCTTGTCCGCCACCACTGAGTTCACCGTGCCTCGTAGAACCG 186
      |||||||::: ||| |||||||
Db 122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
      |||||||::: ||| |||||||
Qy 187 CCGCGAAATGTCCCGAACCCTAGAAAAGCCACCGAGTGGGGTGAAGGAATCGCATATTA 246
      |||||
Db 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
      |||||
Qy 247 TTTAAAGAGATATCATGTCATATAATTTAAAGTACGCTTTATTTATAAATATCATTT 306
      |||||||::: ||| |||||||
Db 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
      |||||||
Qy 307 CAGACGACGACATGGAGCGGACGACATATAGACAGATCAGTATCATCGATATACAGATAGG 366
      |||||
Db 182 MetThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
      |||||
Qy 367 ACGCCGTTTCCATTCAGAGATCAGGATCTAATCGCGGCAAGAGAGATGCTCATCT 426
      |||||
Db 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeuSer 221
      |||||
Qy 427 AAAGCAAGATACCTTAGAACCAATGTATATGTTGAAGCGTTTGACAGGATCGGAGAA 486
      |||||
Db 222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro 241
      |||||
Qy 487 AAACAGTACTTCTAAACCACTCAAAATTCACACGCCCGCAATCTAG----- 534
      ::::::::::|||
Db 242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTrpHisThr 261
      |||||
Qy 534 ----- 534
      ::::::::::|||
Db 262 TyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGly 281
      ::::::::::|||
Qy 534 ----- 534
      ::::::::::|||
Db 282 SerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGlu 301
      ::::::::::|||
Qy 534 ----- 534
      ::::::::::|||
Db 302 IleThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsn 321
      ::::::::::|||
Qy 534 ----- 534
      ::::::::::|||
Db 322 AsnHisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAla 341
      ::::::::::|||
Qy 535 -----GCATGGCACGACTAATGAGAGTATACCGNG 567
      |||||||
Db 342 SerLysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnGluThrTyrThrLys 361
      |||||||
Qy 568 TGGGATCACCATTGATATATCGAAGGAACTCCGTCATTTGATAGTAGGAAATG 627
      |||||
Db 362 IleGlyAlaAlaGlyPheHisHisSerGlyThrSerValAsnCysIleValGluGluVal 381
      |||||
Qy 628 GATCCCGCTCTGTGTTTCCGTATTCATATTTTGGCAATGGCC 669
      |||||

```

```

Db      382 AspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 395
|||||
RESULT 11
US-08-413-118-4
; Sequence 4, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-413-118-4

Alignment Scores:
Pred. No.:      5,97e-65      Length:      1041
Score:          614.50      Matches:    129
Percent Similarity: 44.92%      Conservative: 39
Best Local Similarity: 34.49%      Mismatches: 53
Query Match:      51.21%      Indels:     153
DB:              1          Gaps:         4

US-09-147-052-3_COPY_1_672 (1-672) x US-08-413-118-4 (1-1041)
QY      7 TATTTAGCGGAATGTCATATT----- 30
|||||
Db      22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyLeuAlaAlaThrGlySerArg 41
31 -----TTCCTTAPAGTT 42
42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
|||||
QY      43 ATCTATAT----- 51
|||||
Db      62 ValLeuPheLeuValGlyProArgProValGluGlnSerGlySerThrSerGluGln 81
QY      52 -----GGTACGAACATCTCTCCGACTACCAA 78
|||||

```

```

Db      82 ProArgArgThrValAlaThrProGluValGlyGlyThrProProLysProThrThrAsp 101
QY      79 AATGTGACATCAAGAGAAGTTGTCGAGCGTCCAGTGTCTCTGAGGAAGAG----- 129
|||||
Db      102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
QY      130 ---TCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTGCCTCTAGAACCG 186
|||||
Db      122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
QY      187 CGCGAAATGTCCCGAAGCTAGAAAAGCCAGGAGTGGGTGGGTAAGGAATCGCGATATTA 246
|||||
Db      142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPhethrGluGlyLeuAlaValIle 161
QY      247 TTTAAAGAGATATCAGTCCATATAAATTTAAAGTCAGCGTTTATTATAAAATATCAT 306
|||||
Db      162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrLysAsnIleIle 181
QY      307 CAGACGACGACATGGAGCGGACGACATATAGACACATCACTAATCGATATACAGTAGG 366
|||||
Db      182 MetThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
QY      367 AGCCCGTTTCCATTGAAGAGATCAGGATCTAATCGCGGCAAGGAAGATGCTCATCT 426
|||||
Db      202 ValProValLysValGlnGluIleThrAspLeuLeuAspArgGlyMetCysLeuSer 221
QY      427 AAAGCAAGATACCTTAGAACAAATGATATGTTGAAGCGTTTCACAGGGATGCGGGAGAA 486
|||||
Db      222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgGluAspPro 241
QY      487 AAACAAGTACTTCTAAACCATCAAAATTCACACGCCGCCCAATCTAGG----- 534
|||||
Db      242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyThrHisThr 261
QY      534 ----- 534
Db      262 TyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGly 281
QY      534 ----- 534
Db      282 SerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGlu 301
QY      534 ----- 534
Db      302 IleThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsn 321
QY      534 ----- 534
Db      322 AsnHisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAla 341
QY      535 -----GCATGGCACACGACTAATGAGACGTATACCGTG 567
|||||
Db      342 SerLysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnGluThrTyrThrLys 361
QY      568 TGGGGATACCATGGATATATCGAACGGGACCTCCGTCATTTCTATAGTAGGAGGAATC 627
|||||
Db      362 IleGlyAlaAlaGlyPheHisHisSerGlyThrSerValAsnCysIleValGluGluVal 381
QY      628 GATGCCCGCTCTGTTTCGTTTCGTTTCATATTTCGCAATGCC 669
|||||
Db      382 AspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 395
|||||
RESULT 12
US-08-473-446-4
; Sequence 4, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 STREET: 530 FIFTH AVENUE, 25TH FLOOR
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,446
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/413,118
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: FROMMER, WILLIAM S.
 REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454310-2670
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1041 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: N-terminal
 US-08-473-446-4

Alignment Scores:
 Pred. No.: 5,97e-65 Length: 1041
 Score: 614.50 Matches: 129
 Percent Similarity: 44.92% Conservativeness: 39
 Best Local Similarity: 34.49% Mismatches: 53
 Query Match: 51.21% Indels: 153
 DB: 3 Gaps: 4
 US-09-147-052-3_COPY_1_672 (1-672) x US-08-473-446-4 (1-1041)

QY 7 TATTTAGCGGAATTCGATATT----- 30
 DB 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
 QY 31 -----TTCCTTATAGTT 42
 DB 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
 QY 43 ATTCATAT----- 51
 DB 62 ValLeuPheLeuValGlyProArgProValGluGlnSerGlySerThrSerGluGln 81
 QY 52 -----GGTACGAACATCATCTCCGAGTACCCAA 78
 DB 82 ProArgArgThrValAlaThrProGluValGlyGlyThrProProLysProThrThrAsp 101
 QY 79 AATGTGCATCAGACAGAAAGTTGTTTCGAGCGTCCAGTGTCTGAGGAAGAG----- 129
 DB 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
 QY 130 ---TCTACGTTTATCTTCTCCGCCACAGTGGTTCACCGTGATCCCTCTAGAACCG 186
 DB 122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
 QY 187 CCGCGAAATGTCCTCCCAACCTAGAAAAGCCAGCGAGTGGGTGAAGGAATCGCATTA 246
 DB 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161

QY 247 TTTAAAGAGATATCAGTCCATATAAATTAAGTCACGCTTATTATATAAATATCAT 306
 DB 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
 QY 307 CAGACGACGACATGGACGGGACGACATATAGACATCACTAATCGATATACAGTAGG 366
 DB 182 MetThrThrValTrpSerGlySerTyrAlaValThrAsnArgTyrThrAspArg 201
 QY 367 ACGCCGTTTCCATTGAAGATCAGGATCAATTAATCGACGGCAAGAGAGATGCTCATCT 426
 DB 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSer 221
 QY 427 AAAGCAAGATACCTTAGAAACAATATATATGTTGAAGCGTTTGACAGGATCGCGGAGAA 486
 DB 222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro 241
 QY 487 AAACAAGTACTCTAAACCATCAAAATTCACACGCCGCAATCTAGG----- 534
 DB 242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTrpHisThr 261
 QY 534 ----- 534
 DB 262 TyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGly 281
 QY 534 ----- 534
 DB 282 SerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGlu 301
 QY 534 ----- 534
 DB 302 IleThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsn 321
 QY 534 ----- 534
 DB 322 AsnHisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAla 341
 QY 535 -----GCATGCGACACGACTAATGAGACGTATACCGTG 567
 DB 342 SerLysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnGluThrTyrThrLys 361
 QY 568 TGGGATCACCATGGATATATCGAACGGGACCTCCGTCATTTGTATATAGAGGAATG 627
 DB 362 IleGlyAlaAlaGlyPheHisSerGlyThrSerValAsnCysIleValGluGluVal 381
 QY 628 GATGCCGCTCTGTGTTCCGTATCATATTTGCAATGGCC 669
 DB 382 AspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 395

RESULT 13

US-09-232-468A-2
 ; Sequence 2, Application US/09232468A
 ; Patent No. 6207165
 ; GENERAL INFORMATION:
 ; APPLICANT: AUDONNET et al.
 ; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST PORCINE
 ; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY PATHOLOGIES
 ; FILE REFERENCE: 454313-2230
 ; CURRENT APPLICATION NUMBER: US/09/232,468A
 ; CURRENT FILING DATE: 1999-01-05
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 913
 ; TYPE: PRT
 ; ORGANISM: Pseudorabies virus
 ; US-09-232-468A-2

Alignment Scores:

Score: 2,57e-64 Length: 913
 Percent Similarity: 609.00 Matches: 113
 Best Local Similarity: 73.08% Conservativeness: 39
 Query Match: 54.33% Mismatches: 52
 Indels: 4

```
DB: 3 Gaps: 2
US-09-147-052-3_COPY_1_672 (1-672) x US-09-232-468A-2 (1-913)
QY 52 GGTAGAACTCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAAGTGTGTCG 105
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
QY 106 ACCGTCAGTCTCGAGGAAGAGTCTACGTTTTATCTTTTGTCCGCCACCGAGTGTCA 165
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
QY 166 ACCGTGATCCGCTAGAACCGCGGAAATGTCCGGAACCTAGAGAAACCCACCGAGTGG 225
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 133 ThrValValArgLeuGluProGluAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
QY 226 GGTGAAGGAATCCGATATTTAAAGAGATATCATCTCCATATATAATTTAAAGTGCAG 285
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 153 ThrGluGlyLeuAlaLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
QY 286 CTTTATTATAAAATATCATCTCAGCAGCAGATGGACGGGACGACATATAGACAGATC 345
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 173 IleTyrTyrLysAsnValIleValThrValTrpSerGlySerThrTyrAlaAlaIle 192
QY 346 ACTAATCGATACAGATAGGACCGCGCTTCCATTGAAGATCATCTATGTTCAAGCG 405
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
QY 406 GCGAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACAATGTATGTTCAAGCG 465
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
QY 466 TTTGACAGGGATGCGGGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCC 525
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
QY 526 GAATCTAGGCATGCACACGACTAATGAGACGTATACCGTGTGGGATCCACCATGGATA 585
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 253 GlyThrArgAlaTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
QY 586 TATCGAACGGGAACCTCCGTCATTTGATAGAGAGAAATGGATGCCGCTCTGTGTTT 645
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 273 TyrGlnThrGlyThrSerValAsnCysIleValGluGluValAlaAlaArgSerValTyr 292
QY 646 CCGTATTCTATATTTGCAATGGCC 669
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 293 ProTyrAspSerPheAlaLeuSer 300

RESULT 14
US-09-784-984B-49
; Sequence 49, Application US/09784984B
; Patent No. 6576243
; GENERAL INFORMATION:
; APPLICANT: Meril Ltd.
; APPLICANT: Audonnet, Jean-Christophe
; APPLICANT: Bouchardon, Annabelle
; APPLICANT: Baudu, Philippe
; APPLICANT: Riviere, Michael
; TITLE OF INVENTION: Polynucleotide Vaccine Formula Against Porcine Reproductive and
; FILE REFERENCE: 454313-2230.1
; CURRENT FILING DATE: 2001-02-16
; PRIOR FILING DATE: 1996-07-19
; PRIOR FILING DATE: 1997-07-15
; PRIOR FILING DATE: 1997-07-15
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 913

DB: 3 Gaps: 2
US-09-147-052-3_COPY_1_672 (1-672) x US-09-784-984B-49 (1-913)
QY 52 GGTAGAACTCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAAGTGTGTCG 105
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
QY 106 ACCGTCAGTCTCGAGGAAGAGTCTACGTTTTATCTTTTGTCCGCCACCGAGTGTCA 165
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
QY 166 ACCGTGATCCGCTAGAACCGCGGAAATGTCCGGAACCTAGAGAAACCCACCGAGTGG 225
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 133 ThrValValArgLeuGluProGluAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
QY 226 GGTGAAGGAATCCGATATTTAAAGAGATATCATCTCCATATATAATTTAAAGTGCAG 285
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 153 ThrGluGlyLeuAlaLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
QY 286 CTTTATTATAAAATATCATCTCAGCAGCAGATGGACGGGACGACATATAGACAGATC 345
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 173 IleTyrTyrLysAsnValIleValThrValTrpSerGlySerThrTyrAlaAlaIle 192
QY 346 ACTAATCGATACAGATAGGACCGCGCTTCCATTGAAGATCATCTATGTTCAAGCG 405
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
QY 406 GCGAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACAATGTATGTTCAAGCG 465
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
QY 466 TTTGACAGGGATGCGGGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCC 525
   ::  ::  :::::  :::::  :::::  :::::  |||
Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
QY 526 GAATCTAGGCATGCACACGACTAATGAGACGTATACCGTGTGGGATCCACCATGGATA 585
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 253 GlyThrArgAlaTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
QY 586 TATCGAACGGGAACCTCCGTCATTTGATAGAGAGAAATGGATGCCGCTCTGTGTTT 645
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 273 TyrGlnThrGlyThrSerValAsnCysIleValGluGluValAlaAlaArgSerValTyr 292
QY 646 CCGTATTCTATATTTGCAATGGCC 669
   |||  ::  :::::  :::::  :::::  :::::  |||
Db 293 ProTyrAspSerPheAlaLeuSer 300

RESULT 15
US-08-220-151-7
; Sequence 7, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF INVENTION: CANINE HERPESVIRUS gb, 9C AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
```

```

; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-7

```

```

Alignment Scores:
Pred. No.: 5.75e-64 Length: 868
Score: 606.00 Matches: 107
Percent Similarity: 74.09% Conservative: 36
Best Local Similarity: 55.44% Mismatches: 50
Query Match: 50.50% Indels: 0
DB: 1 Gaps: 0

```

US-09-147-052-3_COPY_1_672 (1-672) x US-08-220-151-7 (1-868)

```

QY 91 AGAAGTTGTTTCGAGCGTCTGTCGAGGAGAGTCTAGCTTTTATCTTTGTCCTCC 150
Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
QY 151 CCACGAGTGGTTCACCGGTGATCCGCTAGAACCGCCGCGGAAATGTCGCCAACCTAGA 210
Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
QY 211 AAAGCCACGAGTGGGTGAGGAGTCCGATATATTATTAAGAGATATCAGTCCATAT 270
Db 81 LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaAlaTyr 100
QY 271 AAATTTAAAGTACGCTTTATTATAAAATATCATTCAGACGACGACATGACGGGACG 330
Db 101 LysPheLysAlaThrValTyrLysAspValIleValSerThrAlaIlePalaGlySer 120
QY 331 ACATATAGACAGATCACTAATCGATATACAGATAGGACGCCGCTTCCATTGAAGAGATC 390
Db 121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
QY 391 ACGGATCTAATCGACGCAAGGAGATGCTCATCTAAAGCAAGATACCTTAGAAACAAT 450
Db 141 ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn 160
QY 451 GTATATGTTGAACGCTTTGACAGGGATGCGGGGAGAAAACAAGTCTTCTAAACCATCA 510
Db 161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProIleAlaSer 180
QY 511 AAATTCACACGCCCCAATCTAGGGCATGGCACGACTAATGAGACGTATACCGGTGG 570
Db 181 LysTyrAsnSerValGlySerLysAlaThrPheHisThrAsnAspThrTyrMetValAla 200
QY 571 GGATCACCATGGATATATCGAAGGACCTCCGCAATGTATAGTAGGAATGGAT 630
Db 571

```

```

Db 201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
QY 631 GCCCGCTCTGTGTTCCGTTATTCATATTTTGGCAATGGCC 669
Db 221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSer 233

```

Search completed: October 8, 2003, 18:21:40
Job time : 19.5802 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:32:42 ; Search time 37.8133 Seconds
(without alignments)
5641.626 Million cell updates/sec

Title: us-09-147-052-3_copy_1_672

Perfect score: 1200
Sequence: 1 atgcacattatttagcgga.....catatttgcattgccaat 672

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p_model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09147052/runat_08102003_154340_29792/app_query.fasta_1.5980
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPGL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=pt -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pt -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052 -CGN_1_1_474 -runat_08102003_154340_29792 -NCPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	100.0	865	14	AAR30169
2	1200	100.0	1086	18	AAW36051
3	659.5	55.0	879	16	AAW77024
4	609	50.7	913	19	AAW68404
5	606	50.5	854	8	AAW70347
6	606	50.5	858	13	AAW22615
7	605	50.4	845	6	AAW50035
8	583	48.6	928	16	AAW77399
9	583	48.6	943	21	AAW32470
10	583	48.6	943	22	AAW50113
11	582	48.5	933	13	AAW27807
12	582	48.5	933	14	AAW41343
13	564.5	47.0	932	19	AAW44947
14	513	42.8	891	17	AAW92746
15	513	42.8	891	19	AAW70293
16	505	42.1	694	12	AAW14666
17	505	42.1	904	12	AAW14665
18	505	42.1	973	12	AAW14680
19	504	42.0	400	8	AAW70654
20	504	42.0	885	17	AAW92747
21	504	42.0	903	8	AAW70426
22	504	42.0	903	8	AAW71135
23	504	42.0	904	14	AAW1779
24	504	42.0	904	17	AAW00375
25	504	42.0	904	18	AAW34553
26	504	42.0	904	22	AAW74441
27	504	42.0	905	9	AAW80915
28	500	41.7	854	19	AAW72113
29	500	41.7	904	9	AAW80914
30	500	41.7	904	14	AAW1778
31	500	41.7	904	17	AAW00376
32	500	41.7	904	18	AAW34552
33	500	41.7	904	19	AAW72193
34	500	41.7	904	22	AAW74442
35	500	41.7	907	8	AAW71136
36	498	41.5	903	7	AAW60244
37	498	41.5	904	23	AAW17812
38	496	41.3	973	8	AAW70769
39	471	39.2	795	19	AAW72062
40	454.5	37.9	250	22	AAW50114
41	427	35.6	903	6	AAW50312
42	426.5	35.5	825	13	AAW22237
43	417	34.8	873	16	AAW79459
44	314	26.2	456	18	AAW36050
45	225	18.8	845	18	AAW26475

ALIGNMENTS

RESULT 1

AAR30169
ID AAR30169 standard; Protein; 865 AA.
AC AAR30169;
XX
XX
DT 25-MAR-2003 (updated)
DT 07-MAY-1993 (first entry)
XX
XX
DE Marek's Disease Virus glycoprotein B homologue of HSV.
XX
XX
KW Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
KW Herpes Simplex Virus.
XX
OS Marek's Disease Virus.
XX
XX
FN EP520753-A1.
XX

PD 30-DEC-1992.
 XX
 XX 24-JUN-1992; 92EP-0305775.
 XX
 PR 28-JUN-1991; 91US-0722860.
 PR 10-DEC-1991; 91US-0803633.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (USDA) US SEC OF AGRIC.
 XX
 XX Lee LF, Li Y, Nazarian K, Ogawa R, Yanagida N;
 PI WPI; 1993-001546/01.
 DR
 XX Recombinant fowl pox virus contg. Marek's disease virus antigen
 PT gene - used to produce cell-free vaccine against Marek's disease
 PT virus
 XX
 PS Example 2; Page 15-19; 30pp; English.
 XX
 CC The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
 CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-SalI
 CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
 CC from K3 fragment were ligated with EcoRI, SalI digested pUC18. The
 CC sequence of the putative MDV gBh was determined by sequencing a set
 CC of deletion mutants. The nucleotide and amino acid sequences were
 CC found to be identical with the published sequences of the gBh of
 CC RBIB strain of MDV (Ross et al., J.Gen.Virol., 70:1789-1894, 1988).
 CC A fragment contg. the entire coding region of MDV gBh was inserted
 CC into pN21729R (see AAQ34774-Q34778) to produce a recombinant FPV/MDVgBh
 CC virus for immunising chickens. See also AAQ34780-Q34781.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 865 AA;

Alignment Scores:
 Pred. No.: 2,72e-129 Length: 865
 Score: 1200.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x AAR30169 (1-865)

QY 1 ATGCACATTTTACGCGAATTCATATTTTCCTATATAGTTATCTATATGCTAGCAAC 60
 Db 1 MethiSTyrPheArgArgAsnGlyPhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTCCGAGCGTCCAGTTGCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCCGACAGTGGGTCAACCGTATCCGCTCTA 180
 Db 41 GluGluGluSerThrPheTyrLeuGlyProProValGlySerThrValIleArgLeu 60
 QY 181 GAACCGCCGCGAAATGTCCGAACTAGAAAGCCAGCGAGTGGGTGAAGGATCCGG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGlyIleAla 80
 QY 241 ATATTATTAAAGAGAATATCATGTCATATATAATTTAAAGTACGCTTTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
 QY 301 ATCATTCAGACGACATGAGCGGGGACGACATATAGACAGATCACTAAATCGATATACA 360
 Db 101 IleIleGlnThrThrTrpThrGlyThrTyrArgGlnIleThrAsnArgTyrThr 120
 QY 361 GATAGGCGCGTTCATTGCAAGAGATCAACGATCAATCGACCGCAAGAGGATGC 420
 Db 121 AsparGThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140

QY 421 TCATCTAAAGCAAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTGACAGGGATGCG 480
 Db 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValIleAlaPheAspArgAspAla 160
 QY 481 GGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACACGCCGGAATCTAGGCGATGG 540
 Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTTP 180
 QY 541 CACAGCACTAATGAGACGCTATACCGTGTGGGATCACCATGATGATATATATCAACGGGAACC 600
 Db 181 HistThrThrAsnGluThrTyrThrValTrpGlySerProTyrPheArgThrGlyThr 200
 QY 601 TCCGCAATGTATAGTAGAGAAATGGATGCCCGCTCTGTCTTCCGTTATTCATATTTT 660
 Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
 QY 661 GCAATGGGCAAT 672
 Db 221 AlaMetAlaAsn 224

RESULT 2
 AAW36051
 ID AAW36051 standard; Protein; 1086 AA.
 AC AAW36051;
 XX
 DT 15-JUL-1998 (first entry)
 XX
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 XX
 KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 XX
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Region 1..672
 FT /note= "derived from Marek's disease virus gB protein"
 FT 693..1086
 FT /note= "derived from M. gallisepticum antigen"
 XX
 PN W09736924-A1.
 XX
 PD 09-OCT-1997.
 XX
 PF 28-MAR-1997; 97WO-JP01084.
 XX
 PR 29-MAR-1996; 96JP-0103548.
 XX
 PA (JAPG) NIPPON ZEON KK.
 XX
 PI Saito S, Tsuzaki Y, Yanagida N;
 XX
 DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96596.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;
 Alignment Scores:

Key	Location/Qualifiers
FT Region	725..741
FT FT	/label= Transmembrane_region
FT Region	747..771
FT FT	/label= Transmembrane_region
XX	
FN	WO9526751-A1.
XX	
PD	12-OCT-1995.
XX	
PF	30-MAR-1995; 95WO-US03982.
XX	
PR	29-MAR-1995; 95US-0413118.
PR	30-MAR-1994; 94US-0220151.
XX	
PA	(VIRO-) VIROGENETICS CORP.
XX	
PI	Limbach KJ, Paoletti E;
XX	
DR	WPI; 1995-366131/47.
DR	N-PSDB; AAT01402.
XX	
PT	Nucleic acids encoding canine herpes virus (CHV) gB, gC and gD glycoproteins - also glyco:proteins and vectors, for the immunisation of neonatal puppies and adult dogs against CHV
PT	
XX	
PS	Example 3; Fig 1A-G; 241pp; English.
XX	
CC	A canine herpesvirus (CHV) protein has a predicted amino acid sequence (AAR77024) that shows significant homology with the gB glycoprotein of numerous herpesviruses. It is the product of a gene (see AAT01402) isolated from CHV genomic DNA using a probe contg. the feline herpesvirus gB, gC and gD genes. CGV glycoprotein gC and gD homologues (AAR77025-26) have also been obtd. These glycoproteins, including recombinant glycoproteins expressed from attenuated recombinant virus vectors, e.g. AIVAC recombinant vCP320 (see AAT01406), can be used in antigenic, immunological or vaccine compositions to protect puppies and adult dogs against CHV.
XX	
SQ	Sequence 879 AA;
Alignment Scores:	
Pred. No.:	7,86e-67 Length: 879
Score:	659.50 Matches: 125
Percent Similarity:	68.10% Conservative: 33
Best Local Similarity:	53.88% Mismatches: 57
Query Match:	54.96% Indels: 17
DB:	16 Gaps: 2
US-09-147-052-3_COPY_1_672 (1-672) x AAR77024 (1-879)	
Qy	25 ATATTTTCCCTATAGTTATCTCA----- 48
Db	8 IilePhePheIleTyThrLeuIleIleCysAspProThrThrProGluSerThrIle 27
Qy	49 -----TATGGTACGAACATCATCTCGAGTACCCAAATGTGCATCAAGAA 96
Db	28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspIleArgGlu 47
Qy	97 GTTGTTCGAGCGTCAGTTG---TCTGAGGAAGAGTCTACGTTTATCTTTGTCCCCCA 153
Db	48 IileLeuArgGluSerGluIleGluSerAspThrSerThrPheTyrmCysProPro 67
Qy	154 CCAGTGGTTCACCGTGATCCGTGTAGAACCGCGGAAATATGTCCCAACCTAGAAA 213
Db	68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrlLysLeu 87
Qy	214 GCCACCGAGTGGGTGAAGGAATCCGGATATATTAAAGAGAATATCAGTCCATATAA 273
Db	88 GlyLysAsnPheThrGluGlyIleAlaIlePheLysGluAsnIleSerProTyrlLys 107
Qy	274 TTTAAAGTGACGCTTTATTATATAAAATATCATTTAGACAGCACCACATGGACGGGAGCA 333

Db 108 PheLysAlaAsnIleTyrTyrLysAsnIleIleIleIleThrThrValTrpSerGlySerThr 127
 QY 334 TATAGACATCACTAATCGATATACAGATAGACGCCGCTTCCATTGAAGATCAGC 393
 Db 128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147
 QY 394 GATCTAATCGAGCGCAAGAGATGCTCATCTAAGCAAGATACCTTAGAACATGTA 453
 Db 148 GluLeuIleAspArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr 167
 QY 454 TATGTTGAAGCTTTCACAGGATCGCGGAGAAAACAAAGTACTTCTAAACCAATCAAAA 513
 Db 168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187
 QY 514 TTCAACACGCCGAATAGGCGATGGCACAGCATATAGACGATATACCGTGGGGA 573
 Db 188 PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly 207
 QY 574 TCACCATGGATATATCGAACGGGAACCTCCGTCATTTGTATAGTAGAGGAATGATGCC 633
 Db 208 GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla 227
 QY 634 CGCTCTGTTCCGTATTCATATTTTGCATGGCC 669
 Db 228 ArgSerValTyrProTyrAspSerPheAlaIleSer 239

RESULT 4
 AAW68404

ID AAW68404 standard; Protein; 913 AA.

XX AC AAW68404;

XX DT 28-OCT-1998 (first entry)

XX DE Aujeszky's disease virus glycoprotein gB.

XX KW Multivalent vaccine; pig; pathogen: respiratory disease; SIV; PRRSV; HCV; digestive disease; Aujeszky's disease virus; pseudorabies virus; vaccine; swine herpesvirus 1; swine influenza virus; hog cholera virus; vector; porcine respiratory and reproductive syndrome virus; glycoprotein; SINRV; swine infertility and respiratory syndrome virus; Actinobacillus pleuropneumoniae.

XX OS Pseudorabies virus.

XX PN FR2751224-A1.

XX PD 23-JAN-1998.

XX PF 19-JUL-1996; 96FR-00093338.

XX PR 19-JUL-1996; 96FR-00093338.

XX PA (INMR) RHONE MERIEUX SA.

XX DR WPI; 1998-112824/11.

XX DR N-PSDB; AAV49291.

XX PT Multi-valent polynucleotide vaccines against porcine pathogens - consist of at least 3 plasmids able to express protective antigens from specified viruses

XX PS Example 8; Fig 2; 63pp; French.

XX CC The invention relates to a multivalent vaccine for protecting pigs against several pathogens, especially pathogens associated with respiratory and digestive diseases. The pathogens are especially selected from Aujeszky's disease virus, swine influenza virus (SIV), porcine respiratory and reproductive syndrome virus (PRRSV), hog cholera virus (HCV) and Actinobacillus pleuropneumoniae. The vaccines are preferably composed of polynucleotide sequences encoding 3 antigens, all as part of vectors. This sequence represents the Aujeszky's disease virus strain NIA3 glycoprotein gB. The coding sequence was subcloned

CC into the plasmid pVR1012 to generate plasmid pAB090 for use in the vaccine.
 CC XX
 SQ Sequence 913 AA;
 Alignment Scores:
 Pred. No.: 5,48e-61 Length: 913
 Score: 609.00 Matches: 113
 Percent Similarity: 73.08% Conservative: 39
 Best Local Similarity: 54.33% Mismatches: 52
 Query Match: 50.75% Indels: 4
 DB: 19 Gaps: 2

US-09-147-052-3_COPY_1_672 (1-672) x AAW68404 (1-913)

QY 52 GGTACGAACTCATCTCCGAGTACCCAA-----AATGTCATCAAGAGAAGTTGTTTCG 105
 Db 95 GlyProSerGluAlaProAspGlyGluTyrGlyLysLeuAspAlaArgThrAlaValArg 114
 QY 106 AGCGTCCAGTTGCTCTGAGGAAGAGTCTACGTTTATCTTTCTCCACACGAGTGGTTCA 165
 Db 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
 QY 166 ACCGTGATCGCTCTAGAACCCCGGAAATGTCCGAACTAGAAACCCAGCGAGTGG 225
 Db 133 ThrValValArgLeuGluProGluGluAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
 QY 226 GGTGAAGGAATCGGATATATTAAAGAGAAATCATGTCAGTCCATATAATTTAAAGTACG 285
 Db 153 ThrGluGlyIleAlaLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
 QY 286 CTTTATTATAAAATATCATTCAGACGACGACATGGAGCGGAGCATATAGACAGATC 345
 Db 173 IleTyrTyrLysAsnValIleValThrThrValTrpSerGlySerThrTyrAlaAlaIle 192
 QY 346 ACTAATCGATATACAGATAGACGCCGCTTCCATTGAAGAGATCAGGATCTAATCGAC 405
 Db 193 ThrAsnArgPheThrAspArgValProValProValGluIleThrAspValIleAsp 212
 QY 406 GGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATGATATGTTGAACGC 465
 Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
 QY 466 TTTCACAGGATCGGGGAGAAAACAAAGTACTTCTAAACCATCAAAATTCACACGCC 525
 Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
 QY 526 GAATCTAGGCGATGGCACACGACTAATGACAGCTATACCGTGTGGGATCACCATGGATA 585
 Db 253 GlyThrArgAlaTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaIleGlyPhe 272
 QY 586 TATCGAACGGGAACCTCCGTCATTTGTATAGTAGAGAAATGGATCCCGCTCTGTGTTT 645
 Db 273 TyrGluThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292
 QY 646 CCGTATTCATATTTTCATGGCC 669
 Db 293 ProTyrAspSerPheAlaLeuSer 300

RESULT 5
 AAP70347

ID AAP70347 standard; Protein; 854 AA.

XX AC AAP70347;

XX DT 25-MAR-2003 (updated)

XX DT 03-OCT-2002 (updated)

XX DT 18-MAY-1991 (first entry)

XX DE Varicella-zoster virus (VSV) immunogenic outer surface protein.
 XX KW Varicella-zoster virus; gB gene product; varicella; vaccine.

Score: 606.00 Matches: 107
 Percent Similarity: 74.09% Conservative: 36
 Best Local Similarity: 55.44% Mismatches: 50
 Query Match: 50.50% Indels: 0
 DB: 13 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x AAP22615 (1-868)

QY 91 AGAGAGTGTGTTTCGAGCGTCCAGTGTCTCAGGAGAGTCTACGTTTATCTTTGTCCTC 150
 DB 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
 QY 151 CCACAGTGGTTCACCGTGTCCGCTAGAACCGCGCGGAAATGTCGCCGAACTAGA 210
 DB 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspThrHis 80
 QY 211 AAGCCACCGAGTGGGTGAGGAATCGCGATATATTAAAGAGAATATACAGTCCATAT 270
 DB 81 LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaAlaTyr 100
 QY 271 AAATTAAAGTACGCTTTATATAAAATATCATTCAGACGACGACATGGACGGGACG 330
 DB 101 LysPheLysAlaThrValTyrLysAspValIleValSerThrAlaTTPAlaGlySer 120
 QY 331 ACATATAGACAGATCACTAATCGATACAGATAGGACGCCGTTCCATTGAAGAGATC 390
 DB 121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
 QY 391 ACGGATCAATCGACGGGAAAGGATGCTCATCTAAAGCAAGATACCTTAGAACAAAT 450
 DB 141 ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn 160
 QY 451 GTATATGTGAAGCGTTTGACAGGATGCGGAGCAAAACAACTCTTCTAAACCATCA 510
 DB 161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
 QY 511 AAATTCACACGCCGCGATAGGCGATGCGACACAGCTAATGAGACGTATACCGTGTGG 570
 DB 181 LysTyrAsnSerValGlySerLysAlaTIPHisThrAsnAspThrTyrMetValAla 200
 QY 571 GGATCACCAGTATATCGACGGGAAACCTCCGTCATGTATAGTAGAGGAATGGAT 630
 DB 201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
 QY 631 GCCGCTCTGTGTTCCGATTCATATATTTGCAATGGCC 669
 DB 221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSer 233

RESULT 7

AAP50035
 ID AAP50035 standard; Protein; 845 AA.

AC AAP50035;
 XX
 AC
 XX
 DT 25-MAR-2003 (updated)
 DT 07-SEP-1991 (first entry)
 XX
 XX
 DE N-terminal sequence of the pseudorabies virus (PRV) gII protein.
 XX
 KW Sub-unit vaccine; antigen; immunoreactive determinant.
 XX
 OS Pseudorabies virus (PRV).
 XX
 FH Key Location/Qualifiers
 FT Modified-site 151..153
 FT FT /label= potential glycosylation site
 FT Modified-site 261..263
 FT FT /label= potential glycosylation site
 FT Modified-site 441..443
 FT FT /label= potential glycosylation site
 FT Modified-site 516..518
 FT FT /label= potential glycosylation site
 FT Modified-site 573..575

FT Modified-site /label= potential glycosylation site
 FT 633..635
 FT /label= potential glycosylation site
 FT 697..699
 FT /label= potential glycosylation site
 XX
 PN EP162738-A.
 XX
 PD 27-NOV-1985.
 PF 09-APR-1985; 85EP-0400704.
 XX
 PR 03-APR-1985; 85US-0719773.
 PR 09-APR-1984; 84US-0598073.
 XX
 PA (MOLE-) MOLECULAR GENETICS INC.
 XX
 PI Robbins AK, Watson RJ, Enquist LW;
 XX
 DR WPI; 1985-298083/48.
 DR N-PSDB; AAN50036.
 XX
 PT Prodn. of Pseudorabies virus sub-unit vaccines - useful for
 PT conferring protection against the virus infections and for
 PT diagnosis esp. with sheep sera
 XX
 PS Example; Fig 13; 153pp; English.
 XX
 CC The inventors claim a recombinant vector comprising a DNA sequence
 CC coding for an immunoreactive and antigenic determinant of a
 CC Pseudorabies virus protein, and a polypeptide having an
 CC immunoreactive and antigenic determinant of a Pseudorabies virus
 CC protein. Vaccines contg. the polypeptide are obtd. economically and
 CC in large amts.; for use in conferring protection against
 CC pseudorabies virus.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 845 AA;

Alignment Scores:

Pred. No.: 1.53e-60 Length: 845
 Score: 605.00 Matches: 112
 Percent Similarity: 72.12% Conservative: 38
 Best Local Similarity: 53.85% Mismatches: 54
 Query Match: 50.42% Indels: 4
 DB: 6 Gaps: 2

US-09-147-052-3_COPY_1_672 (1-672) x AAP50035 (1-845)

QY 52 GGTACGACATCTCTCGAGTACCCAA-----AATGTGACATCAAGAGAAGTTGTTTCG 105
 DB 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
 QY 106 AGCGTCCAGTTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCGCCGAACTAGAAC 165
 DB 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
 QY 166 ACCGTGATCGTCTAGAACCGCGCGGAAATGTCGCCGAACTAGAACCGCGAGTGG 225
 DB 133 ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
 QY 226 GGTGAAGGAATCCGATATATTATAAGAGATATATCAGTCCATATAAATTAAGTGACG 285
 DB 153 ThrGluGlyIleAlaValLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
 QY 286 CTTTATTATAAAATATCTTACAGACGACATGCGGCGGAGCAGCATATACAGATC 345
 DB 173 IleTyrTyrLysAsnValIleValThrValThrPserGlySerThrTyrAlaAlaIle 192
 QY 346 ACTAATCGATATACAGATAGGACGCCGTTTCATTAAGAGATCATCGGATCTTAATCGAC 405
 DB 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212

```

QY 406 GGCAAGGAGATGCTCATCTAAGCAAGATACCTTAGAACCAATGTATATGTTGAGCG 465
Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
QY 466 TTTGACAGGATCGGGAGAAAAAAGTACTTCTAAACCAATCAAAATTCACACGCC 525
Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
QY 526 GAATCTAGGCGATGGCACAGCACTAATAGACGCTATACCGGTGGGGATCACCATGGATA 585
Db 253 GlyThrArgGlyTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
QY 586 TATCGAAGGAACTCCGTCATTTGTATAGTAGAGAAATGATGCCCTCTGTGTTT 645
Db 273 TyrHisThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292
QY 646 CCGTATTTCATATTTTGCATGGCC 669
Db 293 ProTyrAspSerPheAlaLeuSer 300

RESULT 8
AAR77399
ID AAR77399 standard; Protein; 928 AA.
XX
AC AAR77399;
XX
XX 25-MAR-2003 (updated)
DT 01-FEB-1996 (first entry)
XX
XX BHV1 gI glycoprotein.
XX
XX Glycoprotein gI; Immunogen; vaccine.
XX
XX Bovine herpesvirus type 1 P8-2 strain.
XX
XX US5462734-A.
XX
XX 31-OCT-1995.
XX
XX 22-MAR-1993; 93US-0035558.
XX
XX 02-NOV-1990; 90US-0607794.
XX
XX 22-MAR-1993; 93US-0035558.
XX
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
XX Israel BA, Letchworth GJ;
XX
XX WPI; 1995-382370/49.
XX
XX N-PSDB; AAQ94352.
XX
XX Vaccination against bovine Herpes virus 1 - with parenteral and
XX mucosal vaccines contg. glycoprotein gI
XX
XX Disclosure; Column 13-19; 15pp; English.
XX
XX Bovine herpes virus 1 (BHV1) genomic DNA sequences given in AAQ94352-
XX 54 encode the gI, gIII and gIV glycoproteins (AAR77399-401),
XX respectively. Expression of these DNA sequences, e.g. in bovine
XX fibroblasts using a bovine papilloma virus vector, allows prodn.
XX of recombinant, correctly glycosylated glycoproteins useful in
XX vaccine formulations.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 928 AA;

Alignment Scores:
Pred. No.: 5,57e-58 Length: 928
Score: 583.00 Matches: 107
Percent Similarity: 68.45% Conservative: 34
Best Local Similarity: 51.94% Mismatches: 65
Query Match: 48.58% Indels: 0

```

```

DB: 16 Gaps: 0
US-09-147-052-3_COPY_1_672 (1-672) x AAR77399 (1-928)
QY 52 GGTACGAACATCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAAGTTGTTTCGAGGTC 111
Db 97 GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116
QY 112 CAGTTGCTCAGGAAGAGTCTACGTTTATCTTTTGTCTCCCAACAGTGGGTTCACCGTG 171
Db 117 GluAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal 136
QY 172 ATCCGCTAGAACCGCGCGCAAAATGTCGCCGAACCTAGAAAACCCACGAGTGGGTCAA 231
Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
QY 232 GAATCGCGATATTATTAAAGAGAAATATCAGTCCATATAAATTTAAAGTGACGCTTAT 291
Db 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaIleIleTyr 176
QY 292 TATAAAATATCATTCAGACGACATGACGGGGACGACATATAGACAGATCAGATCAAT 351
Db 177 TyrLysAsnValIleValThrThrTyrAlaGlySerThrTyrAlaAlaIleThrAsn 196
QY 352 CGATATACATAGGACGCCCGCTTCCATTGAAGAGATCAGGATCTAATCGACGGCAA 411
Db 197 GluTyrThrAspArgValProValGlyMetGlyIleThrAspLeuValAspLysLys 216
QY 412 GGAAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGAC 471
Db 217 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValAlaAlaPheAsp 236
QY 472 AGGGATCGGGAGAAAAAAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCT 531
Db 237 ArgAspAspAspProTrpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 256
QY 532 AGGCGATGGCACAGCACTAATAGACGCTATACCGTGGGGATCACCATGGATATATCGA 591
Db 257 ArgGlyTrpHisThrThrAspAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 276
QY 592 ACGGGAACCTCCGTCATATTGTATAGTAGAGAAATGATGCCCTCTGTGTTTCCGAT 651
Db 277 ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 296
QY 652 TCATATTTTGCATGGCC 669
Db 297 AspSerPheAlaLeuSer 302

RESULT 9
AAY32470
ID AAY32470 standard; Protein; 943 AA.
XX
XX AAY32470;
XX
XX 27-MAR-2000 (first entry)
XX
XX DNA encoding feline herpesvirus-1 glycoprotein B.
XX
XX FHV; glycoprotein B; raccoon poxvirus; vaccine; cat;
XX feline viral rhinotracheitis.
XX
XX Feline herpesvirus type 1.
XX
XX Key Location/Qualifiers
XX Peptide 1..66
XX Domain /note= "signal peptide"
XX Domain 67..757
XX Domain /note= "extracellular domain"
XX Domain 758..827
XX Domain /note= "transmembrane domain"
XX Domain 828..911
XX Domain /note= "cytoplasmic domain"
XX Modified-site 152..154

```

FT Modified-site /note= "Asn is N-glycosylated"
 FT 261..263
 FT /note= "Asn is N-glycosylated"
 FT 364..366
 FT Modified-site /note= "Asn is N-glycosylated"
 FT 406..408
 FT Modified-site /note= "Asn is N-glycosylated"
 FT 526..528
 FT Modified-site /note= "Asn is N-glycosylated"
 FT 610..612
 FT Modified-site /note= "Asn is N-glycosylated"
 FT 643..645
 FT /note= "Asn is N-glycosylated"
 XX

PN US6010703-A.

XX 04-JAN-2000.

XX 14-AUG-1997; 97US-0911321.

XX 26-JUL-1993; 93US-0096183.

XX (UNMS) UNIV MICHIGAN STATE.

XX Spatz SJ, Maes RK;

XX WPI: 2000-115851/10.

DR N-PSDB; AA235417.

XX New recombinant raccoon poxvirus expressing glycoprotein from feline
 FT herpes virus, used in vaccines to protect against feline viral
 FT rhinotracheitis -

XX Claim 3; Column 31-35; 57pp; English.

XX This sequence represents glycoprotein B (gB) of feline herpesvirus
 CC type 1 (FHV-1) C-27 (ATCC VR 636). The sequence was deduced from a
 CC gB coding region identified in a Sali subfragment (see AA235417) of
 CC the FHV-1 genome. The invention provides a recombinant raccoon
 CC poxvirus which expresses a gene that encodes the FHV gD
 CC glycoprotein precursor (see AA23673) and is inserted into the
 CC thymidine kinase (tk) gene of the poxvirus. The recombinant
 CC poxvirus optionally also contains a gene encoding the FHV-1 gB
 CC precursor inserted, in addition to the gD gene, into the tk gene of
 CC the poxvirus. A live vaccine containing the recombinant raccoon
 CC poxvirus is also claimed. The vaccine is used to protect cats
 CC against feline viral rhinotracheitis caused by FHV-1. It is safe,
 CC stable and more immunogenic than known modified live vaccines,
 CC providing better protection against development of clinical signs
 CC and viral shedding.

XX SQ Sequence 943 AA;

Alignment Scores:

Pred. No.:	5,61e-58	Length:	943
Percent:	583.00	Matches:	120
Percent Similarity:	56.07%	Conservative:	37
Best Local Similarity:	42.86%	Mismatches:	63
Query Match:	48.58%	Indels:	60
DB:	21	Gaps:	5

US-09-147-052-3_COPY_1_672 (1-672) x AA32470 (1-943)

QY 7 TATTTAGCGGAATTCATATTT----- 30
 |||||
 Db 22 TyrPheArgGlnArgCysPhePheProSerLeuGlyLeuAlaThrGlySerArg 41
 QY 31 -----TTCTTATAGTT 42
 |||||
 Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
 |||||
 QY 43 ATTCTATAT-----GTCAGCACTCATCTCCG 69
 |||||

Db	62	ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln	81
QY	70	AGTACCCAAATGTGACATCAGAGAAGTT-----	99
Db	82	ProArgArgThrValAlaThrProGluValGlyValHisHisGlnAsnGlnLeuGlnIle	101
QY	100	-----GTTTCGAGCGTCTGCTGTCTGAGGAAGAG-----	129
Db	102	ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro	121
QY	130	TCTAGCTTTATCTTTGTCCTCCACACAGTGGGTTCAACCGTGATCCGCTTAGAACCGCGC	189
Db	122	SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro	141
QY	190	CGAAATGTCCGGAACCTAGAAAGCCACCGAGTGGGTGAGGAATCCGATATATTT	249
Db	142	ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe	161
QY	250	AAAGAGATATCAGTCCATATAAATTTAAAGTGACGCTTTATTATAAAATATCATTCAG	309
Db	162	LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet	181
QY	310	ACGACGACATGCGGGGACGACATATAGACAGATCACTAATCGATATACAGTAGGACG	369
Db	182	ThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArgVal	201
QY	370	CCCGTTTCCATTGAAGAGATCAGCGATCTAATCGACGCCCGGATCTAGGCATGCGACGACT	549
Db	202	ProValLysValGlnGluIleThrAspIleLeuAspArgArgGlyMetCysLeuSerLys	221
QY	430	GCAAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTGACGGGATCGCGGAGAAAAA	489
Db	222	AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg	241
QY	490	CAAGTACTTCTAAACCATCAAAATTCACACGCCCGGATCTAGGCATGCGACGACT	549
Db	242	GluLeuProLeuLysProProSerSerThrLeuSerArgValArgGlyTrpHis---	260
QY	550	AATGAGACGTATACCGTGTGGGATCACCATGATATATCGAAGGGGAACTCCGTCAT	609
Db	261	AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn	280
QY	610	TGTATAGTAGAGAAATGGATCCCGCTCTGTTCCGTATTCATATTTTCAATGCC	669
Db	281	CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer	300
RESULT 10			
AA50113			
ID	AA50113	standard; Protein; 943 AA.	
AC	AA50113;		
XX	21-DEC-2001 (first entry)		
DT	Feline herpesvirus glycoprotein B recombinant antigen PFHVgB943.		
DE	FHV; glycoprotein B; PFHVgB943; antigen; immune status;		
KW	vaccination status; cat.		
KW	Feline herpesvirus.		
OS	WO200166568-A2.		
XX	13-SEP-2001.		
XX	07-MAR-2001; 2001WO-US07251.		
XX	09-MAR-2000; 2000US-0521738.		
XX	(HESK-) HESKA CORP.		
PA	(COLS) UNIV COLORADO STATE RES FOUND.		
XX	Jensen WA, Lappin MR, Rosen DK, Andrews JS;		
PI			

XX WPI; 2001-639000/73.
 DR N-PSDB; AAH27060.
 PT Determining immune status or vaccination status of an animal to e.g.
 PT calicivirus comprises using a recombinant viral antigen -
 XX Claim 4; Page 87-90; 132pp; English.
 XX The present sequence is that of full-length feline herpesvirus
 CC glycoprotein B recombinant antigen PFHVgB943. The recombinant
 CC antigen, fused to an N-terminal His tag, was produced in Escherichia
 CC coli cells transformed by a recombinant vector comprising nucleic
 CC acid PFHVgB2829 (see AAH27060). PFHVgB943 is an example of a
 CC recombinant infectious agent antigen that can be used in the method
 CC of the invention to determine the immune status of an animal.
 CC The method involves contacting a biological specimen of an animal
 CC (cat, dog or horse) with a recombinant antigen, and detecting the
 CC presence or absence of a complex between the recombinant antigen
 CC and an antibody present in the sample. The method determines
 CC whether the animal is protected against disease or should be
 CC vaccinated. Recombinant antigens (see AAM50107-24), nucleic acids
 CC encoding them (see AAH27054-71), methods of producing them, and
 CC assay methods are provided.
 XX Sequence 943 AA;
 SQ

Alignment Scores:
 Pred. No.: 5,61e-58 Length: 943
 Score: 583.00 Matches: 120
 Percent Similarity: 56.07% Conservative: 37
 Best Local Similarity: 42.86% Mismatches: 63
 Query Match: 48.58% Indels: 60
 DB: 22 Gaps: 5

US-09-147-052-3_COPY_1_672 (1-672) x AAM50113 (1-943)

Qy 7 TATTTTGGCGGAATTCGATATT----- 30
 Db 22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
 Qy 31 -----TTCCTTATAGTT 42
 Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
 Qy 43 ATCTATAT-----GGTAGCACTCATCTCCG 69
 Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
 Qy 70 AGTACCCAAATGTCATCATCAAGAGAGTT----- 99
 Db 82 ProArgArgThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
 Qy 100 -----GTTTCGAGCGTCCAGTTGCTGAGGAAGAG----- 129
 Db 102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
 Qy 130 TCTAGTTTATCTTTGTCCTCCACCGAGTGGGTCAACCGGTGATCCGTAGAACCCGCG 189
 Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
 Qy 190 CGAAATGTCCCGAACCTAGAAAGCCAGCGGTGGGTGAAGGATCGCATATTATT 249
 Db 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
 Qy 250 AAAGAGATATCATCTCATATAAATTTAAAGTGACGCTTTATTATATAAATATCATTCAG 309
 Db 162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleMet 181
 Qy 310 ACGACGACATGCGGGGACACATATAGACAGATCACTATCGATATACAGATAGACG 369
 Db 182 ThrThrValTrpSerGlySerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201

Qy 370 CCCCTTCCATTTGAAGAGATCACCGATCTAATTCAGCGCAAGAAAGAGATGCTCATCTATAA 429
 Db 202 ProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSerLys 221
 Qy 430 GCAAGATACCTTAGAACCAATGTATATGTTCAAGCGTTTGACAGGGATCGGGAGAAAAA 489
 Db 222 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241
 Qy 490 CAAGTACTTCTTAAACCATCAAAATTCACACGCGCGAATCTAGGCGATGGCACACGACT 549
 Db 242 GluLeuProLeuLysProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260
 Qy 550 AATGAGACGTATACCGTGTGGGATCACCATGGATATATCATCAACGCGGACCTCCGTCAAT 609
 Db 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn 280
 Qy 610 TGTATAGTAGAGAAATGGATGCCGCTCTGTTCCGTATTCATATTTTGCATATGCC 669
 Db 281 CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300

RESULT 11

AAR27807

ID AAR27807 standard; Protein; 933 AA.

AC AAR27807;

DT 10-MAR-1993 (first entry)

DE Bovine herpes virus type 1 glycoprotein I.

KW BHV-1; vaccine; gI.

OS Bovine herpes virus type 1.

FH Key Location/Qualifiers

FT Peptide 1..505 /note= "gIb peptide"

FT Peptide 506..933 /note= "gIc peptide"

FT Region 767..828 /note= "putative transmembrane sequence"

XX US151267-A.

XX 29-SEP-1992.

XX 15-JUL-1988; 88US-0219939.

XX 15-JUL-1988; 88US-0219939.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Babiuk L, Fitzpatrick D, van den Hurk S, Zamb T;

XX WPI; 1992-348566/42.

XX N-PSDB; AAQ29515.

XX Protecting bovine hosts from bovine herpes virus type 1 infection -

XX using vaccine contg. polypeptide neutralising epitope of GI, GII or

XX GIII BHV-1 glycoprotein

XX Disclosure; Fig 5; 48pp; English.

XX The sequence is that of the BHV-1 glycoprotein gI which comprises the

XX gIa/gIb/gIc complex, the glycoprotein maintains the proper epitopes

XX necessary to fully protect immunised animals from disease. It elicits

XX antibodies that neutralise virus infectivity and induce complement-

XX mediated cell lysis. It can be used in vaccine compns. which are

XX more protective than killed virus and attenuated live-virus vaccines.

XX See also AAR27808 and AAR27809.

XX Sequence 933 AA;

SQ

Alignment Scores:		7.29e-58	Length:	933
Pred. No.:	582.00	Matches:	107	
Percent Similarity:	68.45%	Conservative:	34	
Best Local Similarity:	51.94%	Mismatches:	65	
Query Match:	48.50%	Indels:	0	
DB:	13	Gaps:	0	
US-09-147-052-3_COPY_1_672 (1-672) x AAR27807 (1-933)				
QY	52	GGTAGCACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTGTTTCGAGCGTC	111	
DB	97	GlyAspAspAlaAlaSerProAsnSerThrAspValArgAlaAlaLeuArgLeuAla	116	
QY	112	CAGTTCTGAGGAAGAGCTAGCTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTG	171	
DB	117	GlnAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal	136	
QY	172	ATCCGCTAGAACCCCGGAAATGTCGGAACCTAGAAAGCCACCGAGTGGGTGAA	231	
DB	137	ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu	156	
QY	232	GGATCGCGATATATTTAAAGAGAATATCAGTCATATAAATTTAAAGTGCAGCTTAT	291	
DB	157	GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIleTyr	176	
QY	292	TATATAATATCATTCAGACGACATGGAGCGGACGACATATAGACATACACTAAT	351	
DB	177	TyrLysAsnValIleValThrThrTyrAlaGlySerThrTyrAlaAlaIleThrAsn	196	
QY	352	CGATATACAGATAGGACCGCTTCCATTCAAGAGATCAGGATCTAATCGACGGGAA	411	
DB	197	GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLys	216	
QY	412	GGAAGATGCTCATCTAAAGCAAGATCCTTAAAGCAATGTATATGTTGAAGCGTTGAC	471	
DB	217	TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValAlaAlaPheAsp	236	
QY	472	AGGATCGCGGAGAAACCAAGTACTTCTAACCACATCAAAATTTCAACGCGCCGAATCT	531	
DB	237	ArgAspAspAspProTrpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal	256	
QY	532	AGGGCATGCCACCACTAATGAGCAGTATACCGTGTGGGATCCACATGGATATATCA	591	
DB	257	ArgGlyTrpHisThrThrAspAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg	276	
QY	592	ACGGGAACCTCCGTAATTTGTATAGTAGGAATGGATGCCGCTGTGTGTTCCGAT	651	
DB	277	ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr	296	
QY	652	TCATATTTGCAATGCC	669	
DB	297	AspSerPheAlaLeuSer	302	
RESULT 12				
ID	AAR41343			
XX	XX	AAR41343 standard; Protein; 933 AA.		
AC	AC	AAR41343;		
XX	XX	18-FEB-1994 (first entry)		
DT	DT	Bovine herpesvirus type I gi glycoprotein.		
DE	DE	BHV-1; Bovine herpesvirus; vaccine; neutralising epitope;		
XX	XX	glycoprotein; coat protein; gi; herpes; virus.		
KW	KW	Bovine herpesvirus.		
XX	XX	Key Location/Qualifiers		
OS	OS	505		
FT	FT	/note= g1a precursor protein cleaved to give g1b		
FT	FT	and g1c		
FT	FT			

FT	Region	767..828
XX	CA2057387-A.	/label= Putative transmembrane region.
PN	12-JUN-1993.	
XX	11-DEC-1991;	91CA-2057387.
XX	11-DEC-1991;	91CA-2057387.
XX	(VETE-) VETERINARY INFECTIOUS DISEASE.	
PI	Babiuk L, Fitzpatrick D, Van DER HURK S, Zamb T;	
XX	WPI: 1993-273267/35.	
DR	P-PSDB; AAR41343.	
XX	Recombinant bovine herpes type 1 proteins gI, gIII and gIV - and	
PT	DNA encoding them, for use in vaccines against e.g. shipping	
PT	fever	
XX	Claim 2; Figure 5; 154pp; English.	

The bovine herpesvirus gI glycoprotein (BHV-1) is used in the construction of vectors which are subsequently used to transform host cells. gI produced by these cells can then be used to produce subunit vaccines comprising one or more neutralising epitopes of the gI glycoprotein. The subunit vaccines are used to protect cattle from disease, especially shipping fever which is a complex syndrome which often includes infection by BHV-1. They are substantially more protective than previous killed or live attenuated virus vaccines.

Sequence 933 AA;

Alignment Scores:
 Pred. No.: 7.29e-58 Length: 933
 Score: 582.00 Matches: 107
 Percent Similarity: 68.45% Conservative: 34
 Best Local Similarity: 51.94% Mismatches: 65
 Query Match: 48.50% Indels: 0
 DB: 14 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x AAR41343 (1-933)

QY	52	GGTAGCACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTGTTTCGAGCGTC	111
DB	97	GlyAspAspAlaAlaSerProAsnSerThrAspValArgAlaAlaLeuArgLeuAla	116
QY	112	CAGTTCTGAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTG	171
DB	117	GlnAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal	136
QY	172	ATCCGCTAGAACCCCGGAAATGTCGGAACCTAGAAAGCCACCGAGTGGGTGAA	231
DB	137	ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu	156
QY	232	GGATCGCGATATATTTAAAGAGAATATCAGTCATATAAATTTAAAGTGCAGCTTAT	291
DB	157	GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIleTyr	176
QY	292	TATATAATATCATTCAGACGACATGGAGCGGACGACATATAGACATACACTAAT	351
DB	177	TyrLysAsnValIleValThrThrTyrAlaGlySerThrTyrAlaAlaIleThrAsn	196
QY	352	CGATATACAGATAGGACCGCTTCCATTCAAGAGATCAGGATCTAATCGACGGGAA	411
DB	197	GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLys	216
QY	412	GGAAGATGCTCATCTAAAGCAAGATCCTTAAAGCAATGTATATGTTGAAGCGTTGAC	471
DB	217	TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValAlaAlaPheAsp	236

XX Black D, Eberle R, Hilliard J, Scinicariello F;
XX WPI; 1996-105220/11.
XX N-PSDB; AAT16474.
XX
XX Detection of herpes B virus by PCR amplification of sample DNA - to
XX detect a specific herpes simian monkey B virus DNA segment.
XX
XX Claim 1; Column 19-24; 22pp; English.
XX
XX The herpes simian monkey B virus proteins, such as the gB
XX glycoprotein (UL27), have immense potential use in the development
XX of serological immunoassays, which can specifically detect virus
XX antigens and/or antibodies to B virus. One approach is to
XX synthesize peptides which, based on the properties of the predicted
XX protein sequence, are likely to be immunologically active. Such
XX serum antibodies which recognize this specific peptide sequence.
XX Synthetic peptides may also be used to produce antibodies against
XX specific regions of the gB glycoprotein which are unique to one
XX virus. These can then be used to develop virus-specific
XX immunoassays for differentiation of B virus from other primate
XX alpha-herpes virus and for identification of antibodies directed
XX against B virus in primate serum samples.
XX
XX Sequence 891 AA;

Alignment Scores:
Pred. No.: 6,71e-50 Length: 891
Score: 513.00 Matches: 95
Percent Similarity: 63.59% Conservative: 36
Best Local Similarity: 46.12% Mismatches: 75
Query Match: 42.75% Indels: 0
DB: 17 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x AAR92746 (1-891)

QY 52 GGTACGAACCTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTC 111
DB 72 GlyThrAsnAlaSerValGluAlaGlyHisAlaThrLeuArgGluAsnLeuArgAspIle 91
QY 112 CAGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGTTCACCGTG 171
DB 92 LysAlaLeuAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrVal 111
QY 172 ATCCGTCTAGACCCCGCGAAATGTCCCGAACCTAGAAAGCCACCGAGGGGGTGA 231
DB 112 ValGlnPheGluGlnProArgProCysProArgAlaProHisGlyGlnAsnTyrThrGlu 131
QY 232 GGAATCGCGATATTTAAACAGATATTCAGTCCATATAATTTAAAGTGACGCTTTAT 291
DB 132 GlyIleAlaValIlePheGlyGluAsnIleAlaProTyrLysPheLysAlaThrMetTyr 151
QY 292 TATAAATATCATTCAGACGACGATGGAGGGGACATATAGACATACACTAAT 351
DB 152 TyrLysAspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGly 171
QY 352 CATATACAGATAGACGCGCTTCCATTTGAAGAGATCAGGATCTAATCGACGGCAA 411
DB 172 IlePheGluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaArg 191
QY 412 GGAAGATGCTCATTAAGACAGATACCTAGAACAAATGTATATGTTGACCGTTTGC 471
DB 192 GlyValCysArgSerThrAlaLysTyrValArgAsnMetGluSerThrAlaPheHis 211
QY 472 AGGGATCGGGGAGAAAACAGTACTTCTTAAACCATCAAAATTCACACGCGCCGAATCT 531
DB 212 ArgAspAspGluSerAspMetLysLeuLysProAlaLysAlaAlaThrArgThrSer 231
QY 532 AGGGATCGCACACGACCTAATGAGACGATATACCGGTGGGGATCACCATGATATATCG 591
DB 232 ArgGlyTrpHisThrThrAspLeuLysTyrAsnProSerArgIleGluAlaPheHisArg 251

QY 592 ACGGGAACCTCCGTCAATTGTATAGTAGAGAAATGATGCCGCTCTGTGTTCCGTAT 651
DB 252 TyrGlyThrThrValAsnCysIleValGluValGluAlaArgSerValTyrProTyr 271
QY 652 TCATATTTTGCATATGCC 669
DB 272 AspGluPheValLeuAla 277

RESULT 15

AAW70293
ID AAW70293 standard; Protein; 891 AA.
XX
AC AAW70293;
XX
DT 06-NOV-1998 (first entry)
XX
DE Simian herpesvirus B gB glycoprotein sequence (UL27).
XX
KW Simian herpesvirus B gB glycoprotein; UL27; ICP protein; UL28;
XX differential diagnostic test; immunoassay; antibody.
XX
OS Simian herpesvirus B.
XX
PN US5767265-A.
XX
PD 16-JUN-1998.
XX
PF 10-OCT-1995; 95US-0541878.
XX
PR 01-APR-1993; 93US-0042747.
XX 10-OCT-1995; 95US-0541878.
XX
PA (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.
XX
PI Black D, Eberle R, Hilliard J, Scinicariello F;
XX
DR WPI: 1998-361791/31.
XX N-PSDB; AAV33167.
XX
PT Monkey herpes B virus DNA - coding for gB glycoproteins and
XX polypeptides
XX
PS Claim 2; Columns 19-24; 22pp; English.
XX
CC The invention provides a Simian herpesvirus B DNA sequence coding
XX for the present gB glycoprotein (UL27) and a portion of an ICP
XX 18.5 kDa protein (UL28; AAW70294). The invention uses these DNA and
XX protein sequences as a basis for the development of differential
XX diagnostic tests for the rapid identification of Simian herpesvirus B
XX cases. Therefore, the virus can be detected by detecting the DNA
XX sequence and knowledge of the amino acid sequence will help in the
XX design of DNA probes and of peptides for use in immunoassays and for
XX antibody production.
XX
SQ Sequence 891 AA;

Alignment Scores:

Pred. No.: 6,71e-50 Length: 891
Score: 513.00 Matches: 95
Percent Similarity: 63.59% Conservative: 36
Best Local Similarity: 46.12% Mismatches: 75
Query Match: 42.75% Indels: 0
DB: 19 Gaps: 0

US-09-147-052-3_COPY_1_672 (1-672) x AAW70293 (1-891)

QY 52 GGTACGAACCTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTC 111
DB 72 GlyThrAsnAlaSerValGluAlaGlyHisAlaThrLeuArgGluAsnLeuArgAspIle 91
QY 112 CAGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGTTCACCGTG 171
DB 92 LysAlaLeuAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrVal 111

Db 92 LysAlaLeuAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrVal 111
QY 172 ATCCGCTAGAACCGCGCGGAAAAATGTCGCCGAACCTAGAAAAAGCCACCCGAGTGGGTGAA 231
Db : : : : :
112 ValGlnPheGluGlnProArgProCysProArgAlaProHisGlyGlnAsnTyrThrGlu 131
QY 232 GGAATCGCGATATTATTTAAAGAGATATCACTCCATATATAATTTAAAGTGAGCGTTTAT 291
Db : : : : :
132 GlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyr 151
QY 292 TATAAAATATATCATTCAGACGACGACATGCGGCGACGACATATAGACAGATCACTAAT 351
Db : : : : :
152 TyrLysAspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGly 171
QY 352 CGATATACAGATAGGACGCCGCTTCCATTGAAGAGATCAGGATCAATCGACGGCAAA 411
Db : : : : :
172 IlePheGluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaArg 191
QY 412 GGAAGATGCTCATCTAAAGCAAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTGAC 471
Db : : : : :
192 GlyValCysArgSerThrAlaLysTyrValArgAsnMetGluSerThrAlaPheHis 211
QY 472 AGGATGCGGGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCT 531
Db : : : : :
212 ArgAspAspGluSerAspMetLysLeuLysProAlaLysAlaAlaThrArgThrSer 231
QY 532 AGGGCATGGCACACGACTAATGAGACGTATACCGTGTGGGGATCACCATGGATATATCGA 591
Db : : : : :
232 ArgGlyTrpHisThrThrAspLeuLysTyrAsnProSerArgIleGluAlaPheHisArg 251
QY 592 ACGGGAACCTCGTCAATTGTATAGTAGAGAAATGGATGCCGCTCTGTGTTTCCGTAT 651
Db : : : : :
252 TyrGlyThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 271
QY 652 TCATATTTTGCATGGCC 669
Db : : : : :
272 AspGluPheValLeuAla 277

Search completed: October 8, 2003, 17:53:45
Job time : 45.8133 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:02 ; Search time 284.054 Seconds
(without alignments)
5924.998 Million cell updates/sec

Title: US-09-147-052-3
Perfect score: 5665
Sequence: 1 atgcactatttagcgga.....gaccgcgtacattttataa 3261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09147052/runat_08102003_154341_29846/app_query.fasta_1.5980
-DB-SPTREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_771_@runat_08102003_154341_29846 -NCPU=3
-NO_MMALP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -FGAPOPOP=10 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_prodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
1	3598	63.5	865	12	Q8JLW4	Q8JLW4 turkey herp

2	3595.5	63.5	805	12	Q8Y44	Q8Y44 turkey herp
3	3595.5	63.5	865	12	Q8JLW2	Q8JLW2 turkey herp
4	3588.5	63.3	805	12	Q8Y45	Q8Y45 turkey herp
5	3587	63.3	865	12	Q8JLW3	Q8JLW3 turkey herp
6	3585	63.3	865	12	Q8JLW3	Q8JLW3 turkey herp
7	3089	54.5	865	12	Q8PQZ1	Q8PQZ1 turkey herp
8	3074	54.3	865	12	Q8Y406	Q8Y406 turkey herp
9	3054.5	53.9	870	12	Q8PQ9	Q8PQ9 meleagrid h
10	3048.5	53.8	864	12	Q8Y408	Q8Y408 meleagrid h
11	3048.5	53.8	870	12	Q8E1G4	Q8E1G4 meleagrid h
12	1965	34.7	879	12	Q8ISD0	Q8ISD0 canine herp
13	1930	34.1	881	12	Q84735	Q84735 phocine her
14	1909.5	33.7	948	12	Q90050	Q90050 feline herp
15	1805	31.9	919	12	Q9QAP5	Q9QAP5 caprine her
16	1803	31.8	931	12	Q9QAP6	Q9QAP6 rangiferine
17	1785.5	31.5	868	12	Q8Y4M4	Q8Y4M4 human herp
18	1779.5	31.4	980	12	Q66682	Q66682 equine herp
19	1763.5	31.1	950	12	Q66678	Q66678 equine herp
20	1756	31.0	906	12	Q91WV3	Q91WV3 pseudorabie
21	1755.5	31.0	975	12	Q93275	Q93275 equine herp
22	1747.5	30.8	891	12	Q9QAP4	Q9QAP4 cervid herp
23	1747.5	30.8	929	12	Q9QAP4	Q9QAP4 cervid herp
24	1747	30.8	947	12	Q8QV01	Q8QV01 bovine herp
25	1746.5	30.8	912	12	Q8QV01	Q8QV01 buffalo her
26	1743	30.8	943	12	Q86665	Q86665 feline herp
27	1740.5	30.7	891	12	Q66018	Q66018 simian herp
28	1739	30.7	904	12	Q91C63	Q91C63 herpes simp
29	1738	30.7	904	12	Q89920	Q89920 herpes simp
30	1737	30.7	901	12	Q69465	Q69465 herpes simp
31	1735	30.6	904	12	Q37453	Q37453 herpes simp
32	1734	30.6	901	12	Q69464	Q69464 herpes simp
33	1729.5	30.5	920	12	Q8YJ98	Q8YJ98 herpesvirus
34	1721	30.4	902	12	Q69095	Q69095 herpes simp
35	1719.5	30.4	885	12	Q69387	Q69387 cercopithec
36	1712.5	30.2	908	12	Q65538	Q65538 baboon herp
37	1707.5	30.1	893	12	Q65540	Q65540 baboon herp
38	1706.5	30.1	904	12	Q69526	Q69526 human herpe
39	1705.5	30.1	904	12	Q9QIM8	Q9QIM8 human herpe
40	1705	30.1	916	12	Q66016	Q66016 cercopithec
41	1702.5	30.1	903	12	Q69076	Q69076 human herpe
42	1702.5	30.1	904	12	Q91WU4	Q91WU4 human herpe
43	1698.5	30.0	904	12	Q9DXE4	Q9DXE4 herpes simp
44	1667	29.4	894	12	Q9YV92	Q9YV92 macropodid
45	1653	29.2	887	12	Q9YV93	Q9YV93 macropodid

ALIGNMENTS

RESULT 1

Q8JLW4	Q8JLW4	PRELIMINARY;	PRT;	865 AA.
AC	Q8JLW4;			
DT	01-OCT-2002 (Tremblrel. 22, Created)			
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	Glycoprotein B.			
GN	GB.			
OS	Turkey herpesvirus.			
OC	Viruses; dsDNA viruses, no RNA stage: Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=10390;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=X;			
RA	Parcells M.S., Shamblin C.E., Dienglewiecz R.L.;			
RT	"DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three			
RT	Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDVs.;"			
RT	in the Glycoprotein L-encoding Gene in Some vv+MDVs.;"			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AV129965; AAM97698.1;			
DR	InterPro; IPR000234; Glycoprot_B.			
DR	Pfam; PF000606; Glycoprotein_B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			

SQ SEQUENCE 865 AA; 98107 MW; 4D2628B5E4DEB28B CRC64;

Alignment Scores:

Pred. No.: 6.63e-192 Length: 865
Score: 3598.00 Matches: 718
Percent Similarity: 87.81% Conservative: 17
Best Local Similarity: 85.78% Mismatches: 60
Query Match: 63.51% Indels: 42
DB: 12 Gaps: 8

US-09-147-052-3 (1-3261) x Q8JLW4 (1-865)

QY 1 ATGCACTATTAGCGGAATGCATATTTTCCTTAGTATTCTATATGGTACGAAC 60
DB 1 MetHisLysIlePheArgAsnGlnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCCGAGTACCAAAATGTGACATCAAGAGAGTGTTCGAGCGTCAGTGTGCT 120
DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGGAGAGTCTAGCTTTTATCTTTGTCTCCCAACAGTGGTCAACCGTGCCTCTA 180
DB 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAACCGCGCGAAAATGCCGAACCTAGAAAAGCCAGCGAGTGGTGAAGGATCCGC 240
DB 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGlyIleAla 80
QY 241 ATATTATTTAAAGAGATATCAGTCCATATAAATTTAAAGTACGCTTTTATTATAAAAT 300
DB 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCAGACGACGATGACCGGAGCAGCATATAGACAGATCACTAATCGATATACA 360
DB 101 IleleGlnThrThrTrpThrGlyThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGGACGCCGCTTCCATTTGAAGAGATCACGGATCTAATCGACGCGCAAGGAAGTGC 420
DB 121 AspArgThrProValSerIleGluGluIleThrAspLeuLeaspGlyLysGlyArgCys 140
QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTGAAGCGTTTGACAGGATCGC 480
DB 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAspAla 160
QY 481 GGAAAAAACAAGTACTCTTAARACCATCAAAATTCACACGCCGCCGATCTAGGCGATGG 540
DB 161 GlyLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACACGACTAATGACAGCTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACC 600
DB 181 HisThrThrAsnGluThrTyrThrValIleTrpGlySerProIleTyrArgThrGlyThr 200
QY 601 TCCGTCAATTTGTATAGTAGAGAAATGGATGCCGCTGTGTGTTCCGTATTCATATTTT 660
DB 201 SerValAsnCysIleValGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGCCAATGGCAGATCGGACATATCTCCATTTTATGTCTATCCCAACAGAG 720
DB 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu 240
QY 721 GCTGCCGAGAACCCATGGGATATCCCAAGGATAATTTCAAACTAGATAGTATTTT 780
DB 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe 260
QY 781 TCAATGGATTGACAGCGTCCAAAAGCAAGCGTTCACAGTCAAGGTAAGTCTTCATC 840
DB 261 SerMetAspLeuAspLysArgLysAlaSerLeuProValLysArgAsnPheLeuIle 280
QY 841 ACATCACATTCACAGTGGGTGGGACTGGCTCCAAAACACTACTGTGTATCTCAATC 900
DB 281 ThrSerHisPheThrValGlyTrpAspThrAlaProLysThrThrArgValCysSerMet 300
QY 901 ACTAAGTGGAAAGAGGTGACTGAATGTTGCGTGCACACAGTTAATGGGAGATACAGATT 960

DB 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
QY 961 ATGGCCCGTGAACCTTCGGCAACGTTTATCAGTATACGACTAGTATTCATCAATCGC 1020
DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
QY 1021 ATCATATTAGGACAAATGTTAAAGCGGAGGAGAGAGCAATCGACAGATATTTAGG 1080
DB 341 IleleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
QY 1081 ACAAAATAATATGACAGTCAAGCTCAAGCTTGACATGTACAAATATTTCTTGCTCTCGG 1140
DB 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
QY 1141 GGATTTATTTAGCATATCAGCCCTGTTCTATCCAAATCCCTGGCTCATATGATCTCAGA 1200
DB 381 GlyPheIleValSerTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
QY 1201 CAATTGATGAGAGACAACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCA 1260
DB 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATAGAAAATGCTACTCTCATTTGTCAGATTGCGCGGAGATATTCGAAATGCACCA 1320
DB 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro 440
QY 1321 AATAGAAAAATTAACATTAGACGACACACACCTTAATTAATCGACATCGTCTGTTCAATTC 1380
DB 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerValGlnPhe 460
QY 1381 GCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATTAATGATGTTTAGTAGG 1440
DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
QY 1441 ATTCGCCACAGCTTGGTGGAAATTCAGATAGAGAATGTTGTTTATGGCAGAGGATA 1500
DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyLe 500
QY 1501 AAGATTAAATCTACGCTACAGCGAGTGCACATTAGAGAGAGAGTGGTGCAACATG 1560
DB 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
QY 1561 TTGGGGAGTGGCTGCTGTATCGAGCTGCCTCTATAGATCGGATCGTCACCTTTG 1620
DB 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
QY 1621 CAAATTTCTATCGAGTTATCACATCCACTAATACATGTTTATAGCCGACCATGGTTCTA 1680
DB 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTCATATGGAGAAAACCAAGGAACATACAGGACAACCTCGGTGAAAAACAACGAGTTG 1740
DB 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTCACACGCTAGAGCTGTAGAGCCATGCTCGGCTAATCATCTAGATATTTCTGTTT 1800
DB 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe 600
QY 1801 GGATCCGTTTATCTTTTGAATACTAATTTTGTAAAGATGGTAGACGTCGCCGAT 1860
DB 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
QY 1861 ATACAGATTGCTASCACATTTGTGAGCTTAATCTAACCTCGCTAGAGATCGGAAAT 1920
DB 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640
QY 1921 TTGCCCTTTATCCGTTTACACAAAAGAGAGTTCGGTGATGTTGGTGATTTGGATTATGCA 1980
DB 641 LeuProLeuSerValTyrThrLysGluLeuArgAspValGlyValLeuAspTyrAla 660
QY 1981 GAAGTAGCTGCCGCGCAATCAACTACATGACTTAATTTTATGACATAACAAGTAATA 2040


```

Db      661 GluValAlaAaArgAsnGlnLeuHisGluLeuLysPheTyAspIleAsnLysValIle 680
QY      2041 GAAGTGGATACAAATTACCG-----GGCTGCAGGAATTCGGCTGTATGCTATT 2091
Db      681 GluValAspThrAsnTyAlaPheMetAsnGlyLeuAlaGluLeu----- 695
QY      2092 ACTAAAAAGATGCAAAACCCAAATAAATAGCCAAACCCCAATTA---GAAGCAGCGGAATG 2148
Db      696 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
QY      2149 GAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCC 2208
Db      709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
QY      2209 AAG-----ATGAAGCTAGTTTATCATCTGCT 2235
Db      729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
QY      2236 TATAGTGAAGCTGAACAGTTAAC-----ATAACCTTAATGCA----- 2274
Db      749 PheLeuAlaTyArgTyTyValAsnLysLysLysSerAsnPrometClyAlaLeuTyPro 768
QY      2275 ---ACATTAGAACAACTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCT 2331
Db      769 MetThrThrGluValLysAlaGlnAlaThrArgGluLeuHisGlyGluSerAsp 788
QY      2332 AATACGGATAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCA----- 2382
Db      789 AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluAlaArgGluMet 806
QY      2383 -----TACAAAGCACTAAACCACTTTAGAACACGCTGCTACTAACCTT 2427
Db      807 IleLysTyMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu 823

```

RESULT 2

```

Q98Y44      PRELIMINARY;      PRT;      805 AA.
AC      Q98Y44
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT      01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE      Glycoprotein B (Fragment).
OS      Turkey herpesvirus.
OC      Viruses; dsDNA viruses, no RNA stage: Herpesviridae;
OC      Alphaherpesvirinae; Marek's disease-like viruses.
OX      NCBI_TaxID=10390;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Davidson I.;
RT      "Marek's disease virus - turkey isolate, gB.";
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF350326; AAK37557.2; -.
DR      InterPro; IPR000234; Glycoprot_B.
DR      Pfam; PF00606; Glycoprotein_B; 1.
DR      ProDom; PD000693; Glycoprot_B; 1.
FT      NON_TER      805
SQ      SEQUENCE      805 AA; 91115 MW; FC8EB090885E15AA CRC64;

```

Alignment Scores:

```

Pred. No.:      9,09e-192      Length:      805
Score:      3595,50      Matches:      711
Percent Similarity:      89.73%      Conservative:      14
Best Local Similarity:      88.00%      Mismatches:      48
Query Match:      63.47%      Indels:      35
DB:      12      Gaps:      6

```

US-09-147-052-3 (1-3261) x Q98Y44 (1-805)

```

QY      1 ATGCAGTATTTAGCGCGAATTCATATTTTCCATTATAGTTATTTCTATATGGTACGAAC 60
Db      1 MetHisTyPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20
QY      61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAAGTTGTTTCGAGCGCTCCAGTTGCT 120

```

Db 381 GlyPheIleValAlaIaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
 QY 1201 GAATTGATGAGACAACAGACCGATGAGATGCTCGACCTGTTAAACAATAAGCATGCA 1260
 Db 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnLysHisAla 420
 QY 1261 ATTTATAAGAAAATGCTACCTCATGTCACGATTCGGCGAGATATTCGAAATGACCA 1320
 Db 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgAspIleArgAsnAlaPro 440
 QY 1321 AATAGAAAAATACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTTCAATTC 1380
 Db 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerValGlnPhe 460
 QY 1381 GCCATGCTCAATTTCTTTATGATCATATACAAACCCATATTAATGATATGTTTAGTAGG 1440
 Db 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
 QY 1441 ATTGCCACAGCTTGGTGGCAATTCAGAAATAGAGAACTTCTTTTATGGCAGCAGGATA 1500
 Db 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuThrPheGluGlyIle 500
 QY 1501 AAGATTAACTAGCGGTACAGCGAGTGCACATTAAGGAGGAGTGGCTGCCAAGATG 1560
 Db 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
 QY 1561 TTGGGGATGCTCGCTGCTATCGAGCTGACCTGCTATAGATCGGAAATCCGTCACCTTG 1620
 Db 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
 QY 1621 CAAATTTCTATCGAGTTATCATCATCCACTAATACATGTTATAGCGGACCATGGTCTCA 1680
 Db 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
 QY 1681 TTTTCATATGGGAAACCAAGGAACATACAGGGACAACCTCGGTGAAACACACAGAGTTG 1740
 Db 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
 QY 1741 CTTCCAAACGCTAGAGCGCTGAGAGCCATGCTCGGCTAATCACTGATATTTCTGTTT 1800
 Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe 600
 QY 1801 GGATCGGTATGCTTTATTTGAAACTATAATTTTGAAGTGTAGACGTCGCGAT 1860
 Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
 QY 1861 ATACAGATGCTAGCACATTTGTCGAGCTTAATCTAACCTGCTAGAAATCGGAAAT 1920
 Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuGluAspArgGluIle 640
 QY 1921 TTGCTTTATCGTTTACACAAAGAGTTCGCTGATGTTGGTATTTGGATTATGCA 1980
 Db 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla 660
 QY 1981 GAAGTAGTCGGCGGATCACTACATGAACTAAATTTATGACATAAACAAGTAATA 2040
 Db 661 GluValAlaArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
 QY 2041 GAAGTGATACAAATTTACCGC-----GGCTGCAGGAATTCGGCTGTATGCTATT 2091
 Db 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu----- 695
 QY 2092 ACTAAAAAGATGCAAAACCAATATATGCCAAACCCCAATTA--CAAGCAGCGCAATG 2148
 Db 696 -----PheAsnGlyMetGlyValValGlyGlnAlaIleGlyLys 708
 QY 2149 GAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAGACTATGCC 2208
 Db 709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
 QY 2209 AAG-----ATTGAAGCTGATTTATCATCTGCT 2235
 Db 729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748

QY 2236 TATAGTGAAGCTGAAACAGTTAAC-----AATAACCTTAATGCA----- 2274
 Db 749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
 QY 2275 ---ACATTAGACAACACTAAATAATGGCTAAACACTAATTTAGANTCAGCCATCAACCAAGCT 2331
 Db 769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp 788
 QY 2332 AATACGGATAAAACGACTTTTGAT 2355
 Db 789 AspLeuGluArgThrSerIleAsp 796
 RESULT 3
 Q8JLW2 PRELIMINARY; PRT; 865 AA.
 ID Q8JLW2;
 AC Q8JLW2;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_taxid=10390;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=attenuated GA;
 RA Parcells M.S., Shamblin C.E., Dienglewiez R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 RT Pathotypes of Marek's Disease Viruses (vMDV, vVMDV, vv+MDV): Mutations
 RT in the Glycoprotein L-encoding Gene in Some vv+MDVs."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AV129969; AA097702.1;
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98020 MW; F4F526114EA02DE5 CRC64;
 Alignment Scores:
 Pred. No.: 9,13e-192 Length: 865
 Score: 3595.50 Matches: 723
 Percent Similarity: 86.16% Conservative: 18
 Best Local Similarity: 84.07% Mismatches: 68
 Query Match: 63.47% Indels: 51
 DB: 12 Gaps: 9
 US-09-147-052-3 (1-3261) x Q8JLW2 (1-865)
 QY 1 ATGCACTATTATTAGCGGAAATTCATATTTTCCCTTATAGTATTCTATATGATCGTAAC 60
 Db 1 MetHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCTCCGAGTACCACAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTTGCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTACGTTTATCTTTTCTCCGCCACAGTGGTTCACACCTGATCCGCTCA 180
 Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAACCGCGCGAAATATGTCGCCAACCTAGAAAAGCCACCGAGTGGGGTGAAGGAATCGCG 240
 Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
 QY 241 ATATTATTAAAGAGAAATATCAGTCCATATAATTTAAAGTACGCTTTATTATATAAAAT 300
 Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
 QY 301 ATCATTCAGACGACGACATGGGCGGACGACATATAGACATACATCAATCATGATATACA 360
 Db 101 IleIleGlnThrThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120

QY 361 GATAGGACCGCGCTTCCATTGAAGAGATCAGGATCTAATCGACGGCAAGAGATGC 420
DB 121 AsparthrProValSerIleleuGluIlethrAspLeuIleAspGlyLysGlyArgCys 140
QY 421 TCATCTAAAGCAAGATACCTTTAGAAACAATCTATATGTTGAAGCGTTTGACAGGATCGC 480
DB 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAspAla 160
QY 481 GGAGAAAAAAGTACTTCTTAAACCATCAAAATCAACACGCCCGAATCTAGGCGATGG 540
DB 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACAGGACTAATGAGAGCTATACCGTGTGGGGATCACCATCGATATATCGAACGGGAACC 600
DB 181 HisThrThrAsnGlyThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCCGTCAATTGTATAGTAGGAATGGATCGCGCTCTGTGTTCCGTATTTCATATTTT 660
DB 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGGCCAAATGGGACATCGGAACATATCTCCATTTTATGTTCTATCCCCACAGAG 720
DB 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu 240
QY 721 GCTGCCGAGAACCCATGGGATATCCCAGGATAAATTCAAACAACTAGATAGCTATTTT 780
DB 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe 260
QY 781 TCAATGGATTTGGCAACGCTCGAAAGCAAGCCCTCCAGTCAAGCGTAACCTTCTCATC 840
DB 261 SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle 280
QY 841 ACATCACACTTCACAGTTGGTGGGACTGGCTCCAAAACACTACTCGTGTATGTTCAATG 900
DB 281 ThrSerHisPheThrValGlyTyrAspTrpAlaProLysThrThrArgValCysSerMet 300
QY 901 ACTAAGTGAAGAGGTGACTGAATGTTGGCGTCAACAGTAAATGGGAGATACAGATTT 960
DB 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
QY 961 ATGGCCCGTGAACCTTTCGGCAACGTTTATCAGTAAATCAGCTAGTTTGATCCAAATCGC 1020
DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
QY 1021 ATCATATPAGCAATGATTTAAACGGCAGCAGCAGCAGCAATCGACAGATATTTAGG 1080
DB 341 IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
QY 1081 ACAAAATATATGACATCGCTCAAGCTTGGACATGTGACATATTTCTTGGCTCGGG 1140
DB 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
QY 1141 GGATTTATTGTAGCATATCACCTGTTCTATCCAAATCCCTGGCTCATATGTACTCTCAGA 1200
DB 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
QY 1201 GAATTGTAGAGACAACAGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCA 1260
DB 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATAGAAAAATGCTACCTCATTTGTCACGATTTGGCGGAGATATTCGAAATGCACCA 1320
DB 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgAspIleArgAsnAlaPro 440
QY 1321 AATAGAAAAATACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTTCAATTC 1380
DB 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe 460
QY 1381 GCCATGCTCCAAATTTCTTTATGATCATATACAAACCCCATATTAATGATGTTAGTAGG 1440
DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480

QY 1441 ATTGCCACAGCTTGGTGGCAATTCGAGAATAGAGAACTTGTTTTATGGCAGCAAGGGATA 1500
DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle 500
QY 1501 AAGATTAAATCTAGCGCTACAGCGAGTCAACATTAGGAAGAGAGTAGTGGCTCAAGATG 1560
DB 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet 520
QY 1561 TTGGGGGATGTCGTGCTATCGAGCTGCATCTGTATAGATGCGGAATCCGCTCACTTTG 1620
DB 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
QY 1621 CAAATTTCTATCGCGAGTTATCACATCCACTAATACATGTTATAGCCGCACTGTTCTTA 1680
DB 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTCATATGAGAAAAACAAGGAACATACAGGGACAACTCGGTGAAACAACAGATG 1740
DB 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTCCAACGCTAGAGGCTGTAGAGCCATCTCGGCTAATCATCTGTAGATATTTCTGTTT 1800
DB 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe 600
QY 1801 CGATCCGCTTATGCTTTATTTGAAAACATAATATTTTGTAGATGGTAGCCTGCCGAT 1860
DB 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
QY 1861 ATACAGATGTAGCACATTTGTGAGCTTAACTTAACCTCTAGAGATCGGGAATTT 1920
DB 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640
QY 1921 TTGCCTTTTATCCGTTTACAAAAAGAGATGTGCTGATGTTGGTGTATGTTGATTATGCA 1980
DB 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla 660
QY 1981 GAAGTAGTCTCGCCCAATCAACTACATGAACTTAATTTTATGACATAAACAAGTAATA 2040
DB 661 GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
QY 2041 GAAGTGGATACAAATTAACGCG -----GGCTGCAAGGAATTCGGTGTATGCTATT 2091
DB 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu ----- 695
QY 2092 ACTAAAAAAGATGCAAAACCCAAATAATGCGCAAAACCCCAATTA ---GAAGCAGCGGAATG 2148
DB 696 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
QY 2149 GAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAGACTATGCC 2208
DB 709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
QY 2209 AAG -----ATTGAAGCTAGTTTATCATCTGCT 2235
DB 729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
QY 2236 TATAGTGAAGCTGAACACAGTTAAC -----AATAACCTTTAATGCA ----- 2274
DB 749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
QY 2275 ---ACATTAGAACAACTAAAATGGCTAAAACATAATTTAGAATCAGCCATCAACCAAGCT 2331
DB 769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluSerAsp 788
QY 2332 AATACGGATAAACGACTTTTGTATAATGAACACCCCAATTTAGTTGAAGCA ----- 2382
DB 789 AspLeuGluArgThrSerIleAspGluArg -----LysLeuGluGluAlaArgGluMet 806
QY 2383 -----TACAAAGCACTA -----AAAAACCACTTTAGAA 2409
DB 807 IleLysTyrMetAlaLeuValSerAlaGluArgHisGluLysLysLeuArgArgGlu 826
QY 2410 CAACGTGCTACTAACCTTTGAAGGTTTGTTCATCACTGCTTTATAATTCAAATTCGCAATAAT 2469

QY	541	CACACGACTAATGAGACGATATACCGTGTGGGATACCATCGATATATCAACGGGAACC	600
Db	181	HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleLeuArgThrGlyThr	200
QY	601	TCCGTCAATTCTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGTATTCATATTTT	660
Db	201	SerValAsnCysIleValGluGluMetAspAlaLeuArgSerValPheProTyrSerTyrPhe	220
QY	661	GCAATGGCCAATGGCGACATCGCGAACAATATCTCCATTTTATGGTCTATCCCCACAGAG	720
Db	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu	240
QY	721	GCTGCCGAGAACCCATCGGGATATCCCGCAGTAATTTCAAACAACCTAGATACCTATTTT	780
Db	241	AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelGlyLeuAspSerTyrPhe	260
QY	781	TCATGGATTGGCAAGCGTCGAAAAGCAAGCCTTCACGTCAAGCGTACTTCTTCATC	840
Db	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArg***PheLeuIle	280
QY	841	ACATCACACTTCACAGTTGGGTGGGACTGGGCTCCAAAACACTACTCGTGTATGTTCAATG	900
Db	281	ThrSerHisPheThrValGlyTyrPaspTrpAlaProLysThrThrArgValCysSerMet	300
QY	901	ACTAAGTGGAAAGAGGTGACTGAAATGTCGCTGCAACAGTTAATGGGAGATACACATTT	960
Db	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320
QY	961	ATGGCCCGTGAACTTTTCGGCAAGCTTTATCAGTAANTACGACTGAGTTGATCCAAATCGC	1020
Db	321	MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340
QY	1021	ATCATATATGACAACTATTAACGCGAGCGAGACGACGACATCGACGAGATATTTAGG	1080
Db	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluIleGlnIlePheArg	360
QY	1081	ACAAAATATATGACAGCTCAGCTCAAGGTTGGACAGTGTACAATATTTCTTGGCTCTCGGG	1140
Db	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380
QY	1141	GGATTTATTGTAGCATATCAGCTGTTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA	1200
Db	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400
QY	1201	GAATTGATGAGACAAACAGCGACGTAGATGAGCTCGAGCTGGTAAACAATAAGCATGCA	1260
Db	401	GluLeuMetArgaspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420
QY	1261	ATTTATAAGAAAAATGCTACCTCATTTGTCACGATTGCGGCGAGATATCCGAATGCACCA	1320
Db	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440
QY	1321	AATAGAAAATATACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTCCAATTC	1380
Db	441	AsnArgLysIleThrLeuAspAspThrAlaIleLysSerThrSerSerValGlnPhe	460
QY	1381	GCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATTAATCATATGTTTAGTAGG	1440
Db	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480
QY	1441	ATTGCCACAGCTGGTCCGAATTCGAAATAGAGAACTTTGTTTATGGCACCAAGGGATA	1500
Db	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle	500
QY	1501	AAGATTAACTTAGCGCTACAGCGAGTGCACATTAGGAAGGAGAGTGGCTGCCAAGATG	1560
Db	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaAlaLysMet	520
QY	1561	TTGGGGGATGTCCTGCTGTATCGAGCTGCTGCTATAGATCGCGAATCCGTCACATTG	1620
Db	521	LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540
QY	1621	CAAAATCTATGCGAGTTATCATATCCATCAATCATATGTTATAGCGGACCAATTTGTTCTA	1680

```

|||||
541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTCATATGGAGAAACCAAGAAACATACAGGGCAAACTCGGTGAAACAAACAGATTG 1740
Db 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnGlnLeu 580
QY 1741 CTTCCCAACCTAGAGCGCTAGAGCCATCGCTCGGCTAATCATCGTAGATATTTCTGTTT 1800
Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe 600
QY 1801 GGATCCGGTTATGCTTTATTTGAAACTATAATTTTAAAGATGTAGACGTCGCCGAT 1860
Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAsp 620
QY 1861 ATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAACCCCTGCTAGAAGATCGGGAATTT 1920
Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640
QY 1921 TTGCCTTTATCGTTTACACAAAGAGAGTTGCCGTGATGTTGGTGTATGATTATGCA 1980
Db 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla 660
QY 1981 GAAGTAGCTCGCGCAATCACTACATGAACCTTAAATTTTATGACATAACAAAGTAATA 2040
Db 661 GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
QY 2041 GAAGTGGATACAAATTTACGCG-----GGCTCGCAGGAATTCGGCTGTATGCTATT 2091
Db 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu----- 695
QY 2092 ACTAAAAAGATGCAACCAACCAATATATGCGCAACCAATTA---GAAGCAGCGCGAATG 2148
Db 696 -----PheAsnGlyMetGlyValGlnValGlyGlnAlaIleGlyLys 708
QY 2149 GAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATTAAGTTCACACAAAGACTATGCC 2208
Db 709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
QY 2209 AAG-----ATTGAAGCTAGTTTATCATCTGCT 2235
Db 729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleIleAlaGlyLeuValAlaAla 748
QY 2236 TATAGTGAAGCTGAACACTTAC-----AATAACCTTAAATGCA----- 2274
Db 749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
QY 2275 ---ACATTAGAACAACTAAATAATGGCTAAACCTAATTTAGAACTCAGCCATCAACCAAGCT 2331
Db 769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp 788
QY 2332 AATACGGATAAACGACTTTTGAT 2355
Db 789 AspLeuGluArgThrSerIleAsp 796

RESULT 5
Q8JLW3 PRELIMINARY; PRT; 865 AA.
AC Q8JLW3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three

```

```

RT Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations
RT In the Glycoprotein L-encoding gene in Some vv+MDVs.*;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129967; AAM97700.1; -.
DR InterPro: IPR00234; Glycoprot.B.
DR Pfam: PF00606; Glycoprotein.B; 1.
DR ProDom: PD000693; Glycoprot.B; 1.
SQ SEQUENCE 865 AA; 98030 MW; 908B5B4B4E114D CRC64;

Alignment Scores: 2.71e-191 Length: 865
Pred. No.: 3597.00 Matches: 17
Score: 87.69% Conservative: 17
Percent Similarity: 85.66% Mismatches: 61
Best Local Similarity: 63.32% Indels: 42
Query Match: 12 Gaps: 8
DB: 12

US-09-147-052-3 (1-3261) x Q8JLW3 (1-865)

QY 1 ATGCACATTTTATAGCGGAATTCATATTTTCTTATAGTATTCTATATGTGTACGAAC 60
Db 1 MethIstyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTTGTCCCCACCAAGTGGTTCACACCGTGATCCGCTA 180
Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAACCGCGGGAATAATGTCCGAACTAGAAAAGCACCAGTGGGTGAAGGAATCGCG 240
Db 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
QY 241 ATATTATTAAGAGAAATATCAGTCCATATAAATTTAAAGTACGCTTTTATTATAAAT 300
Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTTCAGACGACATGCGGGGACGACATATAGACAGATCATCTAATCGATATACA 360
Db 101 IleIleGlnThrThrThrThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGACGCGCGTTTCCATTGAAGAGATCAGCGATCTAATCAGCGGCAAGGAATGC 420
Db 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
QY 421 TCATCTAAAGCAAGATACCTTAGAACAAATGTATATGTTGAAGCGTTTTCAGACGATGCG 480
Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla 160
QY 481 GGAGAAAAACAAGTACTTCTTAAACCATCAAAATTTCAACACGCGCCGAATCTAGGATGG 540
Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACACACTAATGAGAGCTATACCGTGTGGGGATCACCATTGATATATCGAAGCGGAACC 600
Db 181 HisThrThrAsnGluThrTyrThrValTyrGlySerProTyrIleTyrArgThrGlyThr 200
QY 601 TCCGTCAATGTATAGTAGAGGAATGATGCCGCTCTGCTGTGTTTCCGTTATTCATATTTT 660
Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGGCCCAATGGGACATCGCGAAATATCTCTCATTTTATGTTTATGCTCTATCCACAGAG 720
Db 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProGlu 240
QY 721 GCTGCCGCGAGAACCCATGGGATATCCCGAGGATAAATTTCAACAACCTAGATAGTATTTT 780
Db 241 AlaAlaAlaGluProMetGlyTyrProGlnAsnPheLysGlnLeuAspSerTyrPhe 260
QY 781 TCATATGATTTGGACAGCGCTCGAANAAGCAAGCTTCCAGTCAAGCGTACTTTCTCATC 840

```

Db	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280
QY	841	ACATCACACTTCACAGTTGGGTGGGACGTGGGCTCAAAAACCTACTCGTGTATGTTCAATG	900
Db	281	ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet	300
QY	901	ACTAAGTGGAAAGAGGTGACTGAATGTTGGCGTGCACAGTAAATGGGAGATACAGATTT	960
Db	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320
QY	961	ATGGCCCGTGAACCTTCGCGAACGTTTATCAGTAAATACGACTGAGTTTGATCCAAATCGC	1020
Db	321	MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340
QY	1021	ATCATATTAGGACAAATGATTAAACGGGAGCGAGACAGACACAATCGACGACATATTTAGG	1080
Db	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360
QY	1081	ACAAAATAATATGACAGTCAGCTCAAGGTGGACATGTACAATATTTCTTGGCTCTCGGG	1140
Db	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380
QY	1141	GGATTTATTGTAGCATATCAGCCCTGTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA	1200
Db	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400
QY	1201	CAATTTGATGAGAGACAACAGACCGATGAGATGCTCGACCTGGTAAACATATAGCATGCCA	1260
Db	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420
QY	1261	ATTTATAGAAAATGTACTCTCATTCACGATTGGCGGAGATATTCGAAATGCAACCA	1320
Db	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440
QY	1321	AATAGAAAATAACATTAGACGACACACACAGCTATTAAATCGACATCGTCTGTTCAATTC	1380
Db	441	AsnArgLysIleThrLeuAspAsnThrThrAlaIleLysSerThrSerSerValGlnPhe	460
QY	1381	GCCATGCTCCAATTTCTTTATGATCATATACAACCCATATTAAATGATATGTTTAGTAGG	1440
Db	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480
QY	1441	ATTCGCCACGCTGGTGGCAATTCAGAAATAGAGAACTGTTTTATGGCAGCAAGGGATA	1500
Db	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle	500
QY	1501	AAATTAATCCTAGCGCTACAGCGAGTGCACATTAAGGAAGGAGAGTGGCTGCAAAAGATG	1560
Db	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet	520
QY	1561	TTGGGGATCTGCTGCTGATCGACCTGCACCTGCTATAGATCGGGAATCCGTCACCTTGG	1620
Db	521	LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540
QY	1621	CAAAATTTCTATGCGAGTTATCACATCCACTTAATACATGTTATAGCCGACCATTTGGTTCTA	1680
Db	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu	560
QY	1681	TTTTTCATATGGAGAAAACCAAGGAACATACAGGGACAACCTCGGTGAAAACAACGAGTTG	1740
Db	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu	580
QY	1741	CTTCCAAACGCTACAGCGCTGTAGACCATGCTCGGCTAATCATCVCGTAGATATTTTCGTGTTT	1800
Db	581	LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe	600
QY	1801	GGATCCGGTTATGCTTTATTTGAAAACATAATTTTGTAAAGATGGTAGACGCTCCGAT	1860
Db	601	GlySerGlyTyrAlaAlaPheGluAsnCysAsnPheValLysMetValAspAlaAlaAsp	620
QY	1861	ATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAACCCCTGCTAGAAGATCGGGAATP	1920
Db	621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640

QY	1921	TTGCGCTTTATCGGTTTACACAAAAGAGTGGTGTGATGTTGGTGTATTGGATTATGCA	1980
Db			
QY	641	LeuProLeuSerValtyrThrLysGluLeuArgAspValGlyValLeuAspTyrAla	660
Db			
QY	1981	GAAGTAGTCGCCCGCACTCACTACATGAACCTTAATTTTATGACATAAACAAAGTAATA	2040
Db			
QY	661	GluValAlaAarGargAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle	680
Db			
QY	2041	GAAGTGGATACAAATTAACGCG-----GGGCTGCAGCAATTCGGCTGTATGTCTATT	2091
Db			
QY	681	GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu-----	695
Db			
QY	2092	ACTAAAAAAGATGCAAAACCCAAATAATATGGCAAAACCAATTA---GAAGCAGCGCGAATG	2148
Db			
QY	696	-----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys	708
Db			
QY	2149	GAGTTAACAGATCTAATCAATGCTAAAGCGATGACATAGCTTACCTACACAACTATGCC	2208
Db			
QY	709	ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet	728
Db			
QY	2209	AAG-----ATTGAAGCTAGTTTATCATCTGCT	2235
Db			
QY	729	SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla	748
Db			
QY	2236	TATAGTGAAGCTGAACACAGTTAAC-----AATPACCTTAATGCA-----	2274
Db			
QY	749	PheLeuAlaTyrArgTyrTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro	768
Db			
QY	2275	---ACATTAGAACAACTAAAATATGGCTTAAACCTAATTTAGAAATCAGCCATCAACCAAGCT	2331
Db			
QY	769	MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp	788
Db			
QY	2332	AATACGATAAAACGACTTTGTATATGAACACCCCAAAATTTAGTTGAAGCA-----	2382
Db			
QY	789	AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet	806
Db			
QY	2383	-----TACAAAGCAGCTAAAACCCACTTTAGAACAAACGCTGCTACTACCTT	2427
Db			
QY	807	IleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu	823
Db			
QY	823	291	
QY	83291	PRELIMINARY; PRT; 865 AA.	
AC	Q83291;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Glycoprotein 100 precursor.		
GN	GB.		
OS	Marek disease virus type 1.		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphaherpesvirinae; Marek's disease-like viruses.		
OX	NCBI_TaxID=38013;		
RN	[1]		
RP	SEQUENCE OF 1-498 FROM N.A.		
RC	STRAIN=JM;		
RA	Sousloparov M.A., Bakhtina M.M., Krendelshchikov A.V., Babkin I.V.;		
RT	"PCR-mediated cloning and sequencing of gene encoding the glycoprotein		
RT	complex gp100 gp48 (b-antigen).";		
EL	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; X91988; CAA63039.1; -		
DR	InterPro; IPR000234; Glycoprot_B.		
DR	Pfam; PF00606; Glycoprotein_B; 1.		
DR	ProDom; PD000693; Glycoprot_B; 1.		
KW	Signal.		
FT	SIGNAL		
FT	SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;		
FT	1 21 POTENTIAL.		

Alignment Scores:	
Pred. No.:	3.49e-191
Score:	3585.00
Percent Similarity:	87.57%
Length:	865
Matches:	717
Conservative:	16

Best Local Similarity: 85.66%		Mismatches: 62	
Query Match: 63.28%		Indels: 42	
DB: 12		Gaps: 8	
US-09-147-052-3 (1-3261) x Q83291 (1-865)			
QY	1	ATGCACATNTTTAGGCGGGAATTCGCATATTTTCCCTATAGTTATTTCTATATGTTAGTACGAAC	60
DB	1	MetHisTyrPheArgargAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn	20
QY	61	TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT	120
DB	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40
QY	121	GAGGAAGAGTCTACGTTTATCTTTGTCCTCCCCACAGTGGGTTCACCGTGATCGGTCTA	180
DB	41	GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu	60
QY	181	GRACCCGCCGAAATGTCGCCGAACCTAGAAAGCCACGAGTGGGTGAAGGAATCCGC	240
DB	61	GluProProArgLysCysLeuGluProArgLysAlaThrGluTrpGlyGluGlyIleAla	80
QY	241	ATATTATTTAAAGAGAAATATCAGTCCATATAAATTTAAAGTGACGCTTTATTATAAAAT	300
DB	81	IleLeuPheLysgluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn	100
QY	301	ATCATTCAGACGACATGGAGCGGGGACGACATATAGACAGATCACTAATCGATATACA	360
DB	101	IleIleGlnThrThrTrpThrGlyThrIleTyrArgGlnIleThrAsnArgTyrThr	120
QY	361	GATAGACGCCGTTTCCATTGAGAGATCAGGATCTAATCGACGGCGAAGGAGATGC	420
DB	121	AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys	140
QY	421	TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTCACAGGGATGCG	480
DB	141	SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla	160
QY	481	GGAGAAAACAAGTACTTCTTAAACCATCAAAATTCACACGCCCGGAATCTAGGCGATGG	540
DB	161	GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180
QY	541	CACAGGACTAATGAGACGTATACCGTGGGATCACCATGATATATCGAACGGGAACC	600
DB	181	HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr	200
QY	601	TCCGTCAATGTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGCTATTTCATATTTT	660
DB	201	SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe	220
QY	661	GCAATGCCCAATGGCGACATCGCGAACATATCTCCATTTTATGGTCTATCCCCACAGAG	720
DB	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu	240
QY	721	GTCGCCGACAGACCCATGGGATATCCCGAGGATAATTTCAACAACTAGATAGCTATTTT	780
DB	241	AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe	260
QY	781	TCAATGGATTTGGACAGCGTCCGAAAGCAAGCGCTTCCAGTCAAGCGTAACCTTCTCATC	840
DB	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280
QY	841	ACATCACACTTCACAGTTGGTGGGACTGAAATGTTGCGTGCACACAGTTAATGGGAGATACAGATTT	900
DB	281	ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet	300
QY	901	ACTAAGTGAAGAGGTGACTGAAATGTTGCGTGCACACAGTTAATGGGAGATACAGATTT	960
DB	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320
QY	961	ATGCCCGCTGAAGTTCGCGAACGTTTATCAGTAATACGACTGAGTTTGATCCAAATCGC	1020
DB	321	MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340

QY	1021	ATCATATTATAGGACATGTATTAAACCGCAGGAGAGCAATCGACAGATATTTAGG	1080
DB	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360
QY	1081	ACAAAATATATGACAGTCAGCTCAAGCTTGACATGTACATATTTCTTGGCTCTCGG	1140
DB	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380
QY	1141	GGATTATTATGACATATCAGCTGCTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA	1200
DB	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400
QY	1201	GAATTGATGAGAGACACAGGCCGATGAGATGCTCGACCTGGTGAACAATTAAGCATGCA	1260
DB	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420
QY	1261	ATTTATAAGAAAATGCTTACCTCATTTGTCACGATTGCGCGAGATATTCGAATGCACCA	1320
DB	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440
QY	1321	AATAGAAAATAACATTTAGACGACACACAGCTATTAAATCGACATCGCTCTTCAATTC	1380
DB	441	AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe	460
QY	1381	GCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATTAATGATATGTTTAGTAGG	1440
DB	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480
QY	1441	ATTGCCACAGCTTGTGCGAATTGTCAGATACAGAACTGTTTATGGCACCAGGGATA	1500
DB	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle	500
QY	1501	AAGATTAAATCTAGCGCTACAGCGAGTGCACCATTTAGGAAGGAGAGTGGCTCAAGATG	1560
DB	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet	520
QY	1561	TGGGGGATGTCGCTGCTGATTCGAGTGCACCTGCTATAGATCGGGAATCCCTACTTTG	1620
DB	521	LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540
QY	1621	CAAAATTCATCGGAGTTATCATCCACTAATACATGTATATAGCCGACCATTTGTTCTA	1680
DB	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCystyrSerArgProLeuValLeu	560
QY	1681	TTTTCATATGGAGAAAACCAAGAAACATACAGGCACTCGGTGAAACCAACAGATGG	1740
DB	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu	580
QY	1741	CTTCCAACGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCATCTATATTTCTGTGTT	1800
DB	581	LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe	600
QY	1801	GGATCCGTTATGCTTTTATTTGAAAACATATAATTTTGAAGTGTGACGCTGCCGAT	1860
DB	601	GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp	620
QY	1861	ATACAGATGCTAGCACATTTCTCGAGCTTAATCTAACCCCTCTAGAAGATCGGAAATTT	1920
DB	621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640
QY	1921	TTGCTTTATCCGTTTACAAAGAGAGTTCGCGTGTGTTGGTGTATGTTGATTATGCA	1980
DB	641	LeuProLeuSerValTyrThrLysGluLeuArgAspValGlyValLeuAspTyrAla	660
QY	1981	GAAGTAGTCCGCCGCAATCACTACATGAACCTTAATTTTATGACATAACAAGTAATA	2040
DB	661	GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle	680
QY	2041	GAAGTGGATACAAAATTTACGCG-----GGCTCGAGAAATTCGCGTGTATGCTATT	2091
DB	681	GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu-----	695

QY 2092 ACTAAAAAGATGCAACCAATAATGCGCAACCAACCAATTA---GAAGCAGCGCGAATG 2148
 Db 696 -----PhaSGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
 QY 2149 GAGTTAACAGATCTAATCAATGCTAAAGCATGACATAGCTTTCACATACAGACTATGCC 2208
 Db 709 ValValValGlyAlaAlaGlyAlaIleValSerGlyValSerAlaPheMet 728
 QY 2209 AAG-----ATTTAAGCTAGTCTTATCATCTGCT 2235
 Db 729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
 QY 2236 TATAGTGAAGCTGAACAGTTAAC-----AATAACCTTATGCA----- 2274
 Db 749 PheLeuAlaTyArgTyValAsnLysLeuLysSerAsnProMetLysAlaLeuTyPro 768
 QY 2275 ---ACATTAGAACAACTAAATGGCTTAAACAACTAATTTAGAAATCAGCCATCAACCAAGCT 2331
 Db 769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluSerAsp 788
 QY 2332 AATACGGATAAACAGCTTTTGATAATGAACACCAAAATTTAGTTGAAGCA----- 2382
 Db 789 AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet 806
 QY 2383 -----TACAAAGCACTAAACCACTTTAGAACCAAGCTGCTACTACCTT 2427
 Db 807 IleIysTyMetAlaLeuValSerAlaGluGluArgHisGlyLysLysLeu 823

RESULT 7
 ID Q9PWZ1 PRELIMINARY; PRT; 865 AA.
 AC Q9PWZ1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homolog of HSV-1 major DNA-binding protein (Glycoprotein B).
 GN ORF 35 OR UL27.
 OS Turkey herpesvirus,
 OS Marek's disease virus serotype 2 MDV2, and
 OS Gallid herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390, 36353, 35250;
 [1]
 RN SEQUENCE FROM N.A.
 RC SPECIES-Turkey herpesvirus; STRAIN=HPRS24;
 RA Kato K., Jang H., Izumiya Y., Cai J., Tsushima Y., Miyazawa T.,
 RA Kai C., Mikami T.;
 RT Identification and Transcriptional Analysis of the Marek's Disease
 RT Virus Serotype 2 Genes Homologous to the Glycoprotein B (UL27), the
 RT ICPI8.5 (UL28) and the Major DNA-binding protein (UL29) Genes of
 RT Herpes Simplex Virus Type 1.*;
 RL J. Vet. Med. Sci. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 1; STRAIN=HPRS24;
 RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
 RA Lee Y., Kai C., Takahashi E., Mikami T.;
 RT The complete DNA sequence and transcription map of the unique long
 RT genome region of Marek's disease virus type 2.*;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 3; STRAIN=HPRS24;
 RA Izumiya Y., Jang H., Ono M., Mikami T.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Gallid herpesvirus 3; STRAIN=HPRS24;
 RA Izumiya Y., Jang H., Ono M., Mikami T.;
 RT *A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
 RT Strain HPRS24.*;
 RL Curr. Top. Microbiol. Immunol. 0:0-0(2000).

DR EMBL; AB024711; BAA83751.1; -
 DR EMBL; AB024414; BAA82923.1; -
 DR EMBL; AB049735; BAB16537.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 DR DNA-binding.
 KW SEQUENCE 865 AA; 97309 MW; A20E66C70AA75C35 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 1.17e-163 Length: 865
 Score: 3089.00 Matches: 613
 Percent Similarity: 79.56% Conservative: 80
 Best Local Similarity: 70.38% Mismatches: 134
 Query Match: 54.53% Indels: 44
 DB: 12 Gaps: 10

US-09-147-052-3 (1-3261) x Q9PWZ1 (1-865)

QY 1 ATGCACTATTTAGCGGAATTCGATA---TTTTCTTATAGTATTTATCTATATGTTACG 57
 Db 1 MetAsnHisPheSerGlyIleCysValProLeuPheIleSerValValPheTyrPhePhe 20
 QY 58 AACTCATCTCCGAGTACCCAAATGTGCATCAGAGAACTGTTTCGAGCGTCCAGTTG 117
 Db 21 GlyArgValAlaArgAlaGlnAsnValThrSerArgGluAlaValSerValGlnLeu 40
 QY 118 TCTGAGAGAGAGTCTACGTTTATCTTTGTCCTCCACCTGGTTCACCCGTCGATCCGT 177
 Db 41 SerGluAspGluSerSerPhePheLeuCysProProAlaValGlyThrThrValValArg 60
 QY 178 CTAGACCGCGCGGAAATGTCCTCCGAACTCCAGAACCCAGGAGTGGGTGAAGGAATC 237
 Db 61 LeuGluProProArgLysCysProGluProLeuLysAlaThrGluThrGlyGluGlyLe 80
 QY 238 GCATATATTATTAAGAGAAATATCAGTCCATATAAATTTAAAGTACGCTTTTATATAA 297
 Db 81 AlaIleLeuPheLysGluAsnIleAsnProTyrLysPheLysValThrLeuTyrLys 100
 QY 298 AATATCATTCAGACGACATGCGGCGGACACATATAGACAGATCATCAATCATCAT 357
 Db 101 AsnValIleGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
 QY 358 ACAGATAGGAGCGCCCTTCCATTTGAGAGATCAGGATCATCAATCATCAATCAATCA 417
 Db 121 ThrAspArgThrProValSerIleGluGluIleThrAspValIleAspAlaLysGlyArg 140
 QY 418 TGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTCAAGCGTTTGACAGGAT 477
 Db 141 CysSerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaTyrAspGlyAsp 160
 QY 478 GCGGGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGGAATCTAGGCA 537
 Db 161 AlaGlyGluLysGlnValLeuLeuGlnProSerLysPheAsnThrProGluSerLysAla 180
 QY 538 TGSCACACGACATATGAGACGATATACCGTGTGGGATCACCATGATATATCAACGGGA 597
 Db 181 TrpHisThrThrAsnAspThrTyrThrValTyrGlySerProTyrPyrValTyrArgThrGly 200
 QY 598 ACTCCGTCATATGTATAGTAGAGAAATGGATGCCGCTCTGCTGTTTCGTTATTCATAT 657
 Db 201 ThrSerValAsnCysIleValGluGluMetAspAlaArgAlaAlaPheProTyrSerTyr 220
 QY 658 TTTCGAATGCGCAATGGCGACATCGCGAACAATATCTCCATTTTATGTTTATGTTTATG 717
 Db 221 PheAlaMetAlaAsnGlyAspIleAlaAsnMetSerProPheTyrGlyValAlaProPro 240
 QY 718 GAGGCTGCGCAGAACCCCATGGATATCCCGAGGATAATTTCAACAACATAGATACAT 777
 Db 241 GluAlaAlaAlaGluProMetGlyTyrProLeuGluGlnPheArgGlnIleAspGlyTyr 260
 QY 778 TTTTCAATGGATTGGACAAAGCGTCGAAAGCAAGCCTTCCAGTCAAGCTCAAGCTTCTC 837
 Db 778 TTTTCAATGGATTGGACAAAGCGTCGAAAGCAAGCCTTCCAGTCAAGCTCAAGCTTCTC 837

Db 261 PheProMetAspLeuSerLysArgGlnLysAlaSerIleProValLysArgAsnPheLeu 280
 QY 838 ATCACATCACACTTCACAGTTGGTGGGACTGGGCTCCAAAACACTACTCGTGTATGTTCA 897
 Db 281 ValThrProHisPheThrValGlyTrpAspTrpAlaAlaLysThrAlaArgValCysSer 300
 QY 898 ATGACTTAAGTGAAGACAGCTGACTGAAATGTTGGTGCAACAGTAAATGGGAGATACAGA 957
 Db 301 MetValLysTrpLysAspValThrGluMetLeuArgAlaThrValAsnGlyLysTyrArg 320
 QY 958 TTTATGGCCGCGTAACTTCGCAACGGTTTATCAGTAATACGACTGAGTTTGATCCAAAT 1017
 Db 321 PheMetAlaArgGluLeuSerSerThrPheIleSerAsnThrThrGluPheAspProAsp 340
 QY 1018 CGCATCATATTAGGACAATGTATTAAACCGGAGGAGACGACCAATCGACGAGATATTT 1077
 Db 341 ArgIleArgLeuGlyGlnCysValLysArgAspAlaGluThrIleLysArgIlePhe 360
 QY 1078 AGGACAAAATAATACAGCTCAGCTCAAGTTGGACATGTACAATATTTCTGGCTCTC 1137
 Db 361 AlaGlnLysTyrAsnAspSerHisValLysValGlyLysValGlnTyrPheLeuAlaLeu 380
 QY 1138 GGGGATTTATTGTAGCATATCAGCTGTCTTATCCAAATCCCTGGCTCATATGTACCTC 1197
 Db 381 GlyGlyPheLeuValAlaTyrGlnProValMetSerLysSerLeuAlaHisMetTyrLeu 400
 QY 1198 AGAGAATTATGAGACACACAGGACCGGATGAGATGCTCGACTGTAACCAATAGCAT 1257
 Db 401 ArgGluLeuMetArgAspAsnArgThrAspGluLeuLeuAspLeuValHisAsnLysHis 420
 QY 1258 GCAATTTATAAGAAAAATGCTACCTCATTTGTCACAGTTGCGGAGACATATTCGAATGCA 1317
 Db 421 AlaLeuSerAsnLysThrAsnValSerLeuSerArgLeuArgGluLeuArgAsnAla 440
 QY 1318 CCAAATAGAAAAATATACATATAGACACACACAGCTATTAAATCGACATCGCTGTCTCAA 1377
 Db 441 ProLysGluLysLeuAlaLeuAsnGluGlyAlaThrIleArgSerThrSerValGln 460
 QY 1378 TTCGCCATGCTCCAATTTCTTATGATCATATACAAACCCATTAATGATATGTTAGT 1437
 Db 461 PheAlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSer 480
 QY 1438 AGGATTGCCACAGCTTGGTGCGAATTGCGAGATAGAGAACTTGTTTATGGCAGGAGG 1497
 Db 481 ArgIleAlaThrAlaTrpCysGluLeuGlnAsnLysGluLeuAlaLeuTrpGlnGluGly 500
 QY 1498 ATAAAGATTAATCTTAGCGCTACAGCGAGTCAACATTAGAGAGGAGAGTGGCTGCAAA 1557
 Db 501 MetLysIleAsnProSerAlaIleAlaSerAlaThrLeuGlyArgValAlaAlaLys 520
 QY 1558 ATGTTGGGGATGTCGCTGCTATCGAGCTGCAGCTGCTATAGATCGCGAATCCGTCAC 1617
 Db 521 MetLeuGlyAspValThrAlaValSerThrCysThrGlyIleAspAlaGluSerValThr 540
 QY 1618 TTGCAAAATTCATGCGAGTTATCACATCCACTTAATACATGTTATAGCCGACCATGGTT 1677
 Db 541 LeuGlnAsnSerMetArgValAlaThrSerThrAsnMetCysTyrSerArgProLeuVal 560
 QY 1678 CTATTTTCATATGGAGAAACCAAGAAACATACAGGGACAACCTCGGTGAAACACACGAG 1737
 Db 561 LeuPheSerTyrGlyGluAsnGlnGlyArgIleGlnGlyGlnLeuGlyGluAsnAsnGlu 580
 QY 1738 TTGCTTCCAACTAGAGGCTGTAGAGCCATGCTCGGCTTAATCATCGTAGATATTTCTG 1797
 Db 581 LeuLeuProThrLeuGluAlaValGluProCysThrAlaAsnGlnArgTyrPheLeu 600
 QY 1798 TTTGGATCCGTTATGCTTTTAAACTATAATTTTGAAGATGTAGACGCTGCC 1857
 Db 601 PheGlySerAlaTyrAlaLeuPheGluAspTyrAsnPheValLysMetValGluValAla 620
 QY 1858 GATATACAGATTGCTAGCACATTGTCGAGCTTAATCTAACCTGCTAGAGATCGGAA 1917
 Db 621 AspIleGlnThrAlaSerThrPheValAspLeuAsnLeuThrLeuLeuGluAspArgGlu 640

QY 1918 ATTTTCCCTTTATCCGTTTACACAAAAGAGAGTTCCGCGATGTTGGTGTATTTGGATTAT 1977
 Db 641 IleLeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyr 660
 QY 1978 GCAGAGTAGTACGTCGCGCAATCAACTACATGAACCTTAAATTTTATCACATAAACCAAGTA 2037
 Db 661 AlaGluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysVal 680
 QY 2038 ATAGAAGTCGATACAAATTCACGG-----GGCTCGAGGAATTCGGCTGTATGTCT 2088
 Db 681 IleGluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu----- 696
 QY 2089 ATTACTAAAAAGATGCAACCCAAATATATGCGCAACCCAAATTA---GAAGCAGCGCGA 2145
 Db 697 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGly 708
 QY 2146 ATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCCTACACAGACTAT 2205
 Db 709 LysValValValGlyAlaAlaGlyAlaValAlaThrValSerGlyValSerAlaPhe 728
 QY 2206 GCCAAG-----ATTGAAGCTAGTTTATCATCT 2232
 Db 729 MetSerAsnProPheGlyAlaLeuAlaValGlyLeuIleIleAlaGlyLeuValAla 748
 QY 2233 GCTTATAGTGAAGCTGAAACAGTTAAC-----AATACCTTAAATGCA----- 2274
 Db 749 AlaPheLeuAlaTyrTrpTyrValAsnLysLeuArgGlyAsnProMetLysAlaLeuTyr 768
 QY 2275 -----ACATTAGAACAACTAAATAATGGCTAAATACTAATTTAGAAATGCCCATCAACCA 2328
 Db 769 ProValThrThrGluGluLeuLysSerGlnAlaThrArg---GluLeuHisGlyGluGlu 787
 QY 2329 GCTAATACGGAATAAACGACTTTTGTATGATGAACACACCAAAATTTAGTTGAAGCA----- 2382
 Db 788 ProAspSerGluLysAlaProIleAspLysLys-----LysLeuGlnLysAlaArgGlu 805
 QY 2383 -----TACAAAGCACTAAAAACCACTTTAGAACACAGCTGCTACTAACCTTGAAGGT 2433
 Db 806 MetIleLysTyrMetAlaLeuValSerAlaGluArgHisGluAsnAspLeuArgLys 825
 QY 2434 TTGTCTCAACTGCTTATATCAATTCGAATTCGAATAATTTAGTGGATCATACATAAGCT 2493
 Db 826 LysThrArgGlyThrThrAlaIleLeuSerAsnHisLeuThrLysLeuArgLeuLysSer 845
 QY 2494 AGTAGTTTAACTAAAAACACTAGATGACCTA 2526
 Db 846 SerGlyGlnHisTyrAspArgLeuSerProVal 856
 RESULT 8
 Q69406 PRELIMINARY; PRT; 865 AA.
 AC Q69406;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GB homolog.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SB-1;
 RX MEDLINE=94233711; PubMed=8178437;
 RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
 RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
 RT identification and expression by recombinant fowlpox viruses.";
 RL Virology 200;484-493(1994).
 DR EMBL; U01886; AAA19445.1; -
 DR InterPro; IPR000234; Glycoprot.B.
 DR Pfam; PF00606; Glycoprotain_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.

SQ SEQUENCE 865 AA; 97236 MW; 04D4802FDC2E3948 CRC64;

Alignment Scores:

Pred. No.: 7,98e-163 Length: 865
Score: 3074.00 Matches: 611
Percent Similarity: 79.33% Conservative: 80
Best Local Similarity: 70.15% Mismatches: 136
Query Match: 54.26% Indels: 44
DB: 12 Gaps: 10

US-09-147-052-3 (1-3261) x Q69406 (1-865)

```
QY 1 ATGCACATTTTACGGCAATGTCATA---TTTTTCCTATAGTATTCTATATGGTAGC 57
DB 1 MetAsnHisPheSerGlyIleCysValProLeuPheIleSerValValPheTyrPhePhe 20
58 AACTCATCTCCGAGTACCACCAAAATGTACATCAAGAGAGTGTTCGAGGCGTCAGTTG 117
DB 21 GlyArgValAlaArgAlaGlnAsnValThrSerArgGluAlaValSerSerValGlnLeu 40
QY 118 TCTGAGGAAGAGTCTACGTTTATCTTTGTCTCCCAAGTGGTTCACACCGTGATCGGT 177
DB 41 SerGluAspGluSerSerPhePheLeuCysProAlaValGlyThrThrValValArg 60
QY 178 CTAGAACCGCGCGCAAAATGTCGCAACCTAGAAAAGCCACCGAGTGGGTGAAGGAATC 237
DB 61 LeuGluProArgLysCysProGluProLeuLysAlaThrGluTrpGlyGluGlyIle 80
QY 238 GCGATATTATTAAGAGAATATCAGTCCCATATAAATTAAGTGACGCTTTATATAAA 297
DB 81 AlaIleLeuPheLysGluAsnIleAsnProTyrLysPheLysValThrLeuTyrTyrLys 100
QY 298 AATATCATTCAGACGACATGACCGGGAGCAGACATATACAGACATCACTAATCGATAT 357
DB 101 AsnValIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgPhe 120
QY 358 ACAGATAGGACGCGCTTTCATTGAAGAGATACAGCATCTAATCGACGCGCAAGGAAGA 417
DB 121 ThrAspArgThrProValSerIleGluGluIleThrAspValIleAspAlaLysGlyArg 140
QY 418 TGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTGAAGCGTTTGACAGGAT 477
DB 141 CysSerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaTyrAspGlyAsp 160
QY 478 GCGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCGCCGCAATCTAGGCA 537
DB 161 AlaGlyGluLysGlnValLeuLeuGlnProSerLysPheAsnThrProGluSerLysAla 180
QY 538 TGGCACACGACTAATGAGAGCTATACCGTGTGGGATCACCATGGATATATATCAACGGGA 597
DB 181 TrpHisThrThrAsnAspThrTyrThrValTyrGlySerProTrpValTyrArgThrGly 200
QY 598 ACCTCCGCTCAATGTATAGTAGAGGAATGGATGCCGCTCTGTGTTCGTTATTCATAT 657
DB 201 ThrSerValAsnCysIleValGluMetAspAlaArgAlaAlaLeuGlyTyrSerTyr 220
QY 658 TTTGCAATGCCAATGCGACATCGCAACATATCTCCATTTATGGTCTATCCCAACCA 717
DB 221 PheAlaMetAlaAsnGlyAspIleAlaAsnMetSerProPheTyrGlyValAlaProPro 240
QY 718 GAGGCTGCCGCAAGACCCATGGGATATCCCGAGGATAATTCAACAACCTAGATAGCTAT 777
DB 241 GluAlaAlaAlaGluProMetGlyTyrProLeuGluGlnPheArgGlnIleAspGlyTyr 260
QY 778 TTTTCAATGATTGGACAAAGCGTCGAAAAGCAAGCGCTTCCAGTCAAGCGTAACTTTCTC 837
DB 261 PheProMetAspLeuSerLysArgGlnLysAlaSerIleProValLysArgAsnPheLeu 280
QY 838 ATCACATCACATTCACAGTTGGGTGGGCTGGGCTCCAAAACAACTACTCGTATGTTCA 897
DB 281 ValThrProHisPheThrValGlyTyrAspTrpAlaAlaLysThrAlaArgValCysSer 300
QY 898 ATGACTAAGTGGAAAGAGGTGACTGAAATGTTTCGTCGCAACAGCTTAATGGGAGATACAGA 957
```

```
DB 301 MetValLysTrpLysAspValThrGluMetLeuArgAlaThrValAsnGlyLysTyrArg 320
QY 958 TTTATGGCCCTTGAACTTTCGGCAACGTTTATCATAGTAATACGACTAGTTTGTATCCAAAT 1017
DB 321 PheMetAlaArgGluLeuSerSerThrPheIleSerAsnThrThrGluPheAspProAsp 340
QY 1018 CGCATCATATTAGGACCAATGTATTAAACCGCAGGAGCAAGCAGCAATCGACGACAGATATT 1077
DB 341 ArgIleArgLeuGlyGlnCysValLysArgAspAlaGluThrThrIleLysArgIlePhe 360
QY 1078 AGGACAAAATAATAATGACAGTCAGTCAGTCAAGGTTCGACATGTCACAAATATTCTTTCGCCCTC 1137
DB 361 AlaGlnLysTyrAsnAspSerHisValLysValGlyLysValGlnTyrPheLeuAlaLeu 380
QY 1138 GGGGGATTTATGTAGCATATCAGCCTGTCTTATCCAAATCCCTGGCTCATATATAGCTC 1197
DB 381 GlyGlyPheLeuValAlaTyrGlnProValMetSerLysSerLeuAlaHisMetTyrLeu 400
QY 1198 AGAAGATTGATGAGAGACAACAGACGATGAGATGCTCGACCTGGTAAACAATAAGCAT 1257
DB 401 ArgGluLeuMetArgAspAsnArgThrAspGluLeuLeuAspLeuValHisAsnLysHis 420
QY 1258 GCAATTTTATAAGAAAAATGCTACCTATTGTACGATTGGCGGAGATATTTCGAAATGCA 1317
DB 421 AlaLeuSerAsnLysThrAsnValSerLysSerArgLeuArgArgGluLeuArgAsnAla 440
QY 1318 CCAAAATAGAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTTCAA 1377
DB 441 ProLysGlnLysLeuAlaLeuAsnGluGlyAlaThrIleArgSerThrSerValGln 460
QY 1378 TTCCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATTATATGATATCTTTAGT 1437
DB 461 PheAlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSer 480
QY 1438 AGGATTCACACAGCTTGGTGGCAATTCGAGAATAGAGAACCTGTGTTTATGGCAGCAAGG 1497
DB 481 ArgIleAlaThrAlaTrpCysGluLeuGlnAsnLysGluLeuAlaLeuTrpGlnGluGly 500
QY 1498 ATAAGATTAACTAGCGCTACAGCGAGTGCACATACATAGAGGAGGAGTGGTGCNAAG 1557
DB 501 MetLysIleAsnProSerAlaIleAlaSerAlaThrLeuGlyArgArgValAlaAlaLys 520
QY 1558 ATGTGGGGGATGCTGCTGTATCGAGTGCACCTGCTATAGATAGCGGAATCCGTCACCT 1617
DB 521 MetLeuGlyAspValThrAlaValSerThrCysThrGlyIleAspAlaGluSerValThr 540
QY 1618 TTCCAAAATTCATGGAGTTATTCACATCCACCTAAATACATGTTATAGCCGACCATTTGGTT 1677
DB 541 LeuGlnAsnSerMetArgValAlaThrSerThrAsnMetCysTyrSerArgProLeuVal 560
QY 1678 CTATTTTCATATGGAGAAACCAAGGAACATACAGGACAACTCGGTGNAACAACAGGAG 1737
DB 561 LeuPheSerTyrGlyGluAsnGlnGlyArgIleGlnGlyGlnLeuGlyGluAsnAsnGlu 580
QY 1738 TTGCTTCCAAACGCTAGAGCGCTGTAGAGCCATGCTCGCTAAATCATGCTAGATATTTCTG 1797
DB 581 LeuLeuProThrLeuGluAlaValGluProCysThrAlaAsnGlnArgArgTyrPheLeu 600
QY 1798 TTTGGATCCGGTTATGCTTTATTTGAAACTATAATTTTGTGTTAAGATGGTGTAGCGTGC 1857
DB 601 PheGlySerAlaTyrAlaLeuPheGluAspTyrAsnPheValLysMetValGluValAla 620
QY 1858 GATATACAGATTGCTAGCACATTGTTCGAGCTTAATCTAACCTGCTAGAGATCCGGAA 1917
DB 621 AspIleGlnThrAlaSerThrPheValAspLeuAsnLeuThrLeuLeuGluAspArgGlu 640
QY 1918 ATTTGCCCTTTATCCGTTTACACAAAAGAGAGTTCGCTGATGCTGCTGTTATGGATTAT 1977
DB 641 IleLeuProLeuSerValTyrThrLysGluLeuArgAspValGlyValLeuAspTyr 660
QY 1978 GCAGAAGTAGTCGCGCGCAATCAACTACATGAACTTAAATTTTATGACATAACAAGTA 2037
```

```

Db 661 AlaGluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysVal 680
Qy 2038 ATGAAGTGTGATCAAAATTACGGC-----GGCGTCGAGGAATTCGGCTGTATGCT 2088
Db 681 ILeGluValAlaThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu----- 696
Qy 2089 ATTACTAAAGAAGTCAAAACCAATAATAGGCAACCCCAATTA---GAAGCAGCGCGA 2145
Db 697 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGly 708
Qy 2146 ATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCTTCACTCAAGACATAT 2205
Db 709 LysValValValGlyAlaAlaGlyAlaValAlaIleValSerGlyValSerAlaPhe 728
Qy 2206 GCCAAG-----ATTGAAGCTAGTTTATCATCT 2232
Db 729 MetSerAsnProPheGlyAlaLeuAlaValGlyLeuIleIleAlaGlyLeuValAla 748
Qy 2233 GCTTATAGTAGAAGCTGAACAGCTAAC-----AATAACCTTAATGCA----- 2274
Db 749 AlaPheLeuAlaTyrTrpTyrValAsnLysLeuArgGlyAsnProMetLysAlaLeuTyr 768
Qy 2275 -----ACATTAGAACCACTAAATAATGGCTAAATACTAATTTAGAAATCAGCCATCAACCAA 2328
Db 769 ProValThrThrGluGluLeuLysSerGlnAlaThrArg---GluLeuHisGlyGluGlu 787
Qy 2329 GCTAATACGGATAAACAGCTTTTGTATGATGAATGAACACCAAAATTTAGTTGAAGCA----- 2382
Db 788 ProAspSerGluLysAlaProIleAspLysLys-----LysLeuGlnLysAlaArgGlu 805
Qy 2383 -----TACAAAGCACTAAAAACCACTTTAGAACACCGTGTACTACTACCTTGAAGGT 2433
Db 806 MetIleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluAsnAspLeuArgLys 825
Qy 2434 TTGTCACTCAACTGCTTATACTCAAAATTCGCAATTAATTTAGTGGATCTACATAATAAGCT 2493
Db 826 LysThrArgGlyThrThrAlaIleLeuSerAsnHisLeuThrLysLeuArgLeuLysSer 845
Qy 2494 AGTAGTTTANATACTAAACACTAGATCCACTA 2526
Db 846 SerGlyGlnHisTyrAspArgLeuSerProVal 856

RESULT 9
Q9DPQ9
ID Q9DPQ9 PRELIMINARY; PRT; 870 AA.
AC Q9DPQ9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UL27 virion membrane glycoprotein B.
GN HV7035.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RX MEDLINE=20578232; PubMed=11134310;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
RT "The genome of turkey herpesvirus.";
RL J. Virol. 75:971-978(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF291866; AAG45765.1; -
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein_B.1.
DR ProDom; PD000693; Glycoprot_B.1.
SQ SEQUENCE 870 AA; 98816 MW; F93DIC7036938FB1 CRC64;

```

```

Alignment Scores:
Pred. No.: 9.65e-162 Length: 870
Score: 3054.50 Matches: 602
Percent Similarity: 78.14% Conservative: 88
Best Local Similarity: 68.18% Mismatches: 132
Query Match: 53.92% Indels: 61
DB: 12 Gaps: 11

US-09-147-052-3 (1-3261) x Q9DPQ9 (1-870)

Qy 1 ATGCACATATTTAGCGCGGAATTCATATTTTTCCTTATAGTATTATTCATATATGTTACGAC 60
Db 7 MetLysTyrPheAsnArgSerLeuPheIlePheLeuThrProIleLeuSerIleAlaThr 26
Qy 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGCTCCAGTTCTCT 120
Db 27 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 46
Qy 121 GAGGAAGACTCTACGTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 180
Db 47 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 66
Qy 181 GAACCGCGCGAAATGTCGCGAACCTAGAAAGCCACCGAGTGGGTGAAGGAATCGCG 240
Db 67 GluProArgLysCysProGluProLeuLysSerThrGluTrpGlyGluIleAla 86
Qy 241 ATATTATTTAAAGAGAATATCAGTCCATATAATAATTTAAAGTACGCTTTTATATAAAAT 300
Db 87 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 106
Qy 301 ATCATTGAGCAGGACATGAGCGGGAGCAGATATAGACAGATCATCATATCGATATACATA 360
Db 107 ValIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 126
Qy 361 GATAGGACGCGCTTCCATTGAGAGATCACGATCTAATTCGACGCGCAAGGAAGATGC 420
Db 127 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysGlyCys 146
Qy 421 TCATCTAAAGCAAGATACCTTTAGAAACAATGTATATTTGTAAGCGTTTGACAGGAGTGC 480
Db 147 SerSerLysAlaArgTyrLeuArgAsnValTyrValAspAlaTyrAspArgAspGlu 166
Qy 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGCGATGG 540
Db 167 AsnGluLysGlnValLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTrp 186
Qy 541 CACACGACTAATGAGAGCTATACCGTGTGGGATCACCATGGATATATCGAAGGGGAACC 600
Db 187 HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpValTyrArgThrGlyThr 206
Qy 601 TCCGTCAATTGTATAGTAGAGAAATGGATGCCCGCTCTGTGTTCGTTTCGTTATCATATTT 660
Db 207 SerValAsnCysIleValGluGluMetAspAlaArgSerAlaPheProTyrThrTyrPhe 226
Qy 661 GCAATGCCAATGGCGACATCGCGAACATATCTCCATTTTATGCTCTATCCCCACAGAG 720
Db 227 AlaMetAlaAsnGlyAspIleAlaAsnMetSerProPheTyrGlyThrThrProThrAsp 246
Qy 721 GCTCCCGCAGAACCCATGGGATATCCCGCAGGATATTTCAACAACACTAGATAGCTATTTT 780
Db 247 AlaAlaAlaGluProMetSerTyrProGlnAspArgPheArgGlnPheAspSerTyrPhe 266
Qy 781 TCAATGGATTTGGACAAAGCGTCGAAAGCAAGCCCTCCAGTCAAGCGTAACCTTCTCATC 840
Db 267 ProMetAspLeuAspThrArgArgLysLeuValThrProValLysArgAsnPheLeuVal 286
Qy 841 ACATCAGACTTCACAGTTGGGTGGGACTGGGCTCCAAATACTACTCGTGTATGTTCAATG 900
Db 287 ThrGlnHisPheThrValGlyTrpThrTrpAlaProLysIleSerArgValCysSerMet 306
Qy 901 ACTAAGTGGAAAGAGGTGACTGAATGTTCGCTGCACAGCTTATGGGAGATACAGATTT 960
Db 307 AlaLysTrpLysGluValThrGluMetLeuArgAlaArgValGlyLysLysPheArgPhe 326

```


Percent Similarity: 78.14% Conservative: 90
Best Local Similarity: 67.95% Mismatches: 132
Query Match: 53.81% Indels: 61
DB: 12 Gaps: 11

us-09-147-052-3 (1-3261) x Q69408 (1-864)

```
QY 1 ATGCACATATTTAGCGGAAATGTCATATTTTCCTTATAGTATTATCTCATATGTTGGTACGAAC 60
DB 1 MetLysThrPheAsnArgSerLeuPheIlePheLeuThrProIleLeuSerIleAlaThr 20
QY 61 TCATCTCCGAGTACCCAAAATGTGACATCAAGAGAGTGTTCGACGCTCCAGTGTCT 120
DB 21 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 40
QY 121 GAGGAGAGTCTAGCTTTTATCTTTGTGCCCCACAGTGGGTTTCAACCGTGATCCGCTA 180
DB 41 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 60
QY 181 GAACCGCGGAAAATGTCGGAACCTAGAAAACCCAGCAGTGGGTGAAGGAATCGCG 240
DB 61 GluProArgLysCysProGluProLeuLysSerThrGluTrpGlyGluGlyIleAla 80
QY 241 ATATTATTAAAGAGATATCAGTCCATATAAATTTAAAGTGACGCTTTATTATAAAAT 300
DB 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCAGACGACGATGGACGGGACGACATATAGACAGATCATAATCGATATACA 360
DB 101 ValIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGGACGCCGCTTCATTGAAGACATCAGGATCATCGGATCATCGACGGAAGAGATGC 420
DB 121 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysGlyLysCys 140
QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTGAAGGTTTGACAGGAGTGG 480
DB 141 SerSerLysAlaArgTyrLeuArgAsnValThrValAspAlaTyrAspArgAspGlu 160
QY 481 GGAGAAAACAGTACTTCTTAAACCATCAAAATTCACACGCCGCAATCTAGGCGATGG 540
DB 161 AsnGluLysGlnValLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTyr 180
QY 541 CACAGCATATGAGAGATACCGTGTGGGATCACCATGGATATATCGAACCGGAACC 600
DB 181 HisThrThrAsnGluThrTyrThrValTrpGlySerProTyrValTyrArgThrGlyThr 200
QY 601 TCCGTCATATGTATAGTAGAGAAATGGATGCCGCTGTGTTCGTTTCGTTATCATATTTT 660
DB 201 SerValAsnCysIleValGluLeuMetAspAlaArgSerAlaPheProTyrThrTyrPhe 220
QY 661 GCAATGGCCAATGGGACATCGGACATATCTCCATTTTATGTCCTATCCCCACAGAG 720
DB 221 AlaMetAlaAsnGlyAspIleAlaAsnMetSerProPheTyrGlyThrThrProThrAsp 240
QY 721 GCTGCCGACAGACCCATGGGATATCCCGAGATAATTTCAACACACTAGATAGTATTTT 780
DB 241 AlaAlaAlaGluProMetSerTyrProGlnAspArgPheArgGlnPheAspSerTyrPhe 260
QY 781 TCAATGGATTTGGACAGCGTCGAAAACGCAAGCCTTCAGTCAAGCGTAACTTTCATC 840
DB 261 ProMetAspLeuAspThrArgArgLysLeuValThrProValLysArgAsnPheLeuVal 280
QY 841 ACATCAGCTTCACAGTTGGGTGGGACTGGGCTCCAAAACACTACTCGTGTATGTTCAATG 900
DB 281 ThrGlnHisPheThrValGlyTyrThrTipAlaProLysIleSerArgValCysSerMet 300
QY 901 ACTAAGTGGAAAGAGGTGACTGAATGTTCGCTCCACAGTTAATGGGAGATACAGATTT 960
DB 301 AlaLysThrLysGluValThrGluMetLeuArgAlaArgValGlyLysLysPheArgPhe 320
QY 961 ATGGCCCGTGAACCTTCGGCAACGTTATCAGTAAATACGAGTGAAGTTGATCCAAATCGC 1020
```

```
DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAsnProAspArg 340
QY 1021 ATCATATTAGGACAAATGTATTAAACGCGAGGAGAGCAAGCAATCGAGCAGATATTTAGG 1080
DB 341 IleIleLeuGlyGlnCysValLysArgAspAlaGluAlaGluIleAspLysIlePheLys 360
QY 1081 ACAAAATATATGACAGTCAAGTTCAGGTGGACATGTACAAATATTTCTTGGCTCTCGG 1140
DB 361 ArgLysTyrAsnAspThrHisValLysMetGlyHisValGlnTyrPheLeuThrLeuGly 380
QY 1141 GGATTTATTGTAGCATATCAGCCGTCTTATCCAAATCCCTGGCTCATATGTACTCTCAGA 1200
DB 381 GlyPheValValAlaTyrGlnProIleLeuSerLysSerLeuAlaHisMetCysLeuArg 400
QY 1201 GAATTGATGAGAGACAAACAGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCA 1260
DB 401 GluLeuLeuArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATTAAGAAAATGCTACCTCATTTGTCACGATTTGGCGGAGATATTCGAAATGCACCA 1320
DB 421 IleIleSerAsnAsnAlaThrSerValSerArgLeuArgAspLeuHisArgAlaSer 440
QY 1321 AATAGAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTCAATTC 1380
DB 441 SerGlyLysAlaIleLeuAsnAspSerValIleIleArgSerThrAlaSerValGlnPhe 460
QY 1381 GCCATGTCCTCAATTTCTTTATGATCATATACAAACCCATATTAATGATATTTAGTAGG 1440
DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnAlaHisIleAsnGluMetPheSerArg 480
QY 1441 ATTCCACACGCTTGGTGGAAATTCGAAATAGACAACTGTTGTTTATGGCAGCAGGAGATA 1500
DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnLysGluLeuValLeuTyrArgGluAlaIle 500
QY 1501 AAGATTATCTAGCGCTACAGCGAGTGCACATTAGGAAGAGAGTGGCTGCAAGATG 1560
DB 501 LysIleAsnProSerValThrAlaSerAlaThrLeuGlyLysArgValGlyAlaLysMet 520
QY 1561 TTGGGGGATGTCGTGCTGTATCGAGTGCATCTATAGATCGGAAATCCGTCACCTTTG 1620
DB 521 LeuGlyAspValAlaAlaValSerSerCysIleGluIleAspSerAspSerValThrLeu 540
QY 1621 CAATAATCTATGCCAGTTATCACATCCATCAATACATGTTATACCCGACCATTTGTTCTA 1680
DB 541 GlnAsnSerMetArgValValThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTCATATGAGAGAAAACCAAGGAAACATACAGGAGCAACTCGGTGAAACAAACAGAGTTG 1740
DB 561 PheSerTyrGlyAspArgGlnAspLysIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTCCACGCTAGAGCGTGTAGACCATGCTCGGCTAATCATCTAGATATTTCTGTTT 1800
DB 581 IleProThrLeuGluAlaIleGluProCysSerAlaAsnHisArgTyrPheLeuPhe 600
QY 1801 GGATCCGTTATGCTTTTAAACTATATTTGTTAAGATGGTAGAGCGTCCGAT 1860
DB 601 GlyAspGlyTyrAlaPheTyrGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
QY 1861 ATACAGATTGCTAGCACATTTGTCGAGCTTAATCAACCTGCTAGAGAGATCGGAAAT 1920
DB 621 IleGlnLeuAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640
QY 1921 TTGCCTTTATCCGTTTACAAAAGAGAGTGGTGTGATGTTGGTGTATTTGGATTATGCA 1980
DB 641 LeuProLeuSerValTyrThrLysGluLeuArgAspValGlyValLeuAspTyrAla 660
QY 1981 GAAGTACTCGCCCAATCAACTACATCAACTTAAATTTTATGATACATAAACAAGTAATA 2040
DB 661 GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnArgValMet 680
QY 2041 GAAGTGTATACAAATTTACGCG- - - - -GGGCTGCAGGAAATTCGGCTGTATGCTATT 2091
DB 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu- - - - - 695
```

```

QY 2092 ACTAAAAAGATGCAAAACCAATATATGCGCAACCAATTA---GAAGAGCGCGCAATG 2148
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
696 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
QY 2149 GAGTTAAAGATCTAATCAATGCTAAAGCGATGACATTAAGCTTCACTACAAGACTATGCC 2208
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
709 ValValValGlyAlaAlaGlyAlaIleValSerThrValSerGlyLysAlaPheMet 728
QY 2209 AAG-----ATTGAAGCTAGTTTATCATCTGCT 2235
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleValIleAlaGlyLysAlaAla 748
QY 2236 TATAGTGAAGCTGAACAGCTTAAC-----ATAACTTAAATGCA----- 2274
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
QY 2275 ---ACATTAGAACAACTAAAATGGCTAAACAACTAATTTAGATACGCCATCAACCAAGCT 2331
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
769 MetThrThrGluGluLeuLys-----AspGlnAlaThrArgLysPro 782
QY 2332 ATACGGATAAACGACTTTTGATATGAATGAACCAACCAATTTAGTTGAA----- 2379
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 HisThrAspGlySer-----AspSerGluLeuMetSerIleAspGluArgLysLeuGlu 800
QY 2380 -----GCATACAAAGCAGCTAAAACCACTTTAGAACCAACGCTGCTACT 2421
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
801 AlaAlaArgGluMetIleLysTyrMetAlaLeuValSerAlaGluGluArgHisGlnLys 820
QY 2422 AACCTT-----GAAGTTTGTTCATCAACTGCTTATAATCAAAATTCGCAAT--- 2466
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
821 LysLeuArgLysLysLysArgGlyThrThrAlaIleLeuSerAspHisLeuSerAsnMet 840
QY 2467 -----AATTAGTGCATCTACATAAAGCTAGTAGTTTAACTAAACACTA 2517
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
841 ArgLeuMetAsnGlyHisArgLysTyrAspLysLeuAsnAspThrAspSerGluThrAsp 860
QY 2518 GATCCACTA 2526
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 AspGluIle 863

```

RESULT 11

```

Q9EIG4
ID Q9EIG4 PRELIMINARY; PRT; 870 AA.
AC Q9EIG4;
DT 01-NAR-2001 (TREMBLrel. 16, Created)
DT 01-NAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UL27 glycoprotein gb.
GN UL27.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_Taxid=37108;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=FC126;
RX MEDLINE=21195611; PubMed=11297687;
RA Kingham B.F., Zelnik V., Kopacek J., Majerciak V., Ney E.,
RA Schmidt C.J.;
RT "The genome of herpesvirus of turkeys: comparative analysis with
RT Marek's disease viruses.";
RL J. Gen. Virol. 82:1123-1135(2001).
DR EMBL; AF282130; AAG30067.1; -
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR Prodom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 870 AA; 98827 MW; CE02FB7E7F9EA4A2 CRC64;

```

Alignment Scores:

```

Pred. No.: 2 08e-161 Length: 870
Score: 3048.50 Matches: 600
Percent Similarity: 78.14% Conservative: 90

```

```

Best Local Similarity: 67.95% Mismatches: 132
Query Match: 53.81% Indels: 61
DB: 12 Gaps: 11

```

US-09-147-052-3 (1-3261) x Q9EIG4 (1-870)

```

QY 1 ATGCACTATTTTGGCGGAATTCATATATTTTCCCTATATAGTTATTCATATATGTCAGCAAC 60
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 MetLysTyrPheAsnArgSerLeuPheLeuThrProIleLeuSerIleAlaThr 26
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCAGCGCTCCAGTTGTCT 120
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 SerGluIleLysLeuProAsnValThrAlaArgGluIleValSerGlyIleGlnLeuSer 46
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACACAGTGGTTCACCGTGATCCGTCTA 180
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 GluAspGluThrThrPheTyrValCysProProValGlySerThrIleValArgLeu 66
QY 181 GAACCCCGCGAAATGTCCCGAACTAGAAAGCACCAGCTGGGTGAAGGAATCGCG 240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 GluProProArgLysCysProGluProLeuLysSerThrGluTrpGlyGluGlyIleAla 86
QY 241 ATATTATTAAAGACAATATCAGTCCATATAATTTAAAGTACGCTTTATTTATAAAAT 300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 106
QY 301 ATCATTCAGACGACATGAGCGGGGACGACATATAGACAGATCAGTAAATGATATACA 360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 ValIleGlnThrThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 126
QY 361 GATAGGCGCGCTTCCATTGAAGAGATCAGGATCTAATCGCGCAAGGAAGATGC 420
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 AspArgThrProValSerIleAspGluIleThrAspLeuIleAspGlyLysGlyLysCys 146
QY 421 TCATCTAAGCAAGATACCTTAGAACCAATGTATATGTTGAAGCTTTGACAGGATCGG 480
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147 SerSerLysAlaArgTyrLeuArgAsnValTyrValAspAlaTyrAspArgAspGlu 166
QY 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATCAACAGCCCGCAATCTAGGCGATGG 540
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 AsnGluLysGlnValLeuLeuArgProSerLysPheSerThrAlaGluSerArgAlaTrp 186
QY 541 CACAGCACTAATGAGACGTATACCGTGTGGGATCACCATGATATATCGAACCGGAACC 600
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
187 HisThrThrAsnGluThrTyrThrValTrpGlySerProTyrValTyrArgThrGlyThr 206
QY 601 TCCGTCAATTGTATAGTAGAGAAATGATCCGCTCTGTGTTCGCTATCATATTTT 660
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
207 SerValAsnCysIleValGluMetAspAlaArgSerAlaPheProTyrThrTyrPhe 226
QY 661 GCAATGGCCAATGGCGACATCGCAACATATCTCCATTTATGGTCTATCCCCACAGAG 720
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
227 AlaMetAlaAsnGlyAspIleAlaAsnMetSerProPheTyrGlyThrThrProThrAsp 246
QY 721 GCTCGCGCAGAACCCATGGGATATCCCCAGATATTTTCAACAACTAGATAGTATTTT 780
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
247 AlaAlaAlaGluProMetSerTyrProGlnAspArgPheArgGlnPheAspSerTyrPhe 266
QY 781 TCAATGGATTGGACAGCGTCAAGCAAGCCCTTCCAGTCAAGCGTAACTTCTCTCATC 840
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 ProMetAspLeuAspThrArgArgLysLeuValThrProValLysArgAsnPheLeuVal 286
QY 841 ACATCACACTTCACAGTTGGTGGAGCTGGCTCCAAAACACTACGCTGTATGTTCAATG 900
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
287 ThrGlnHisPheThrValGlyTrpThrTrpAlaProLysIleSerArgValCysSerMet 306
QY 901 ACTAAGTGGAAAGAGTCACTCAATGTTGCGTCCACACAGTTAATGGGAGATACAGATT 960
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 AlaLysTrpLysGluValThrGluMetLeuArgAlaArgValGlyLysLysPheArgPhe 326
QY 961 ATGGCCCGTGAACTTTTCGGCAACCTTTATCATAGTAAATACGACTGAGTTGATCCAAATCGC 1020
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
327 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAsnProAspArg 346

```


DB: 12 Gaps: 10

US-09-147-052-3 (1-3261) x Q91SD0 (1-879)

QY 25 ATATTTTTCCTATAGTATTCTGA----- 48

Db 8 IlePhePheIleIleTyrrThrLeuIleIleCysAspProThrThrProGluSerThrIle 27

QY 49 -----TATGGTACGAACCTCATCTCCGAGTACCCAAATGTGCATCAACGAGAA 96

Db 28 AsnProLeuAsnHisAsnLeuSerThrProLysProThrSerAspAspIleArgGlu 47

QY 97 GTTGTTCGAGCGTCCAGTTG--NCTGAGGAAGAGTCTACGTTTATCTTTTGTCCGCCA 153

Db 48 IleLeuArgGluSerGlnIleGluSerAspThrSerThrPheTyrMetCysProPro 67

QY 154 CCAGTGGGTCAACCGTGTATCGTGTAGAACCGCGGAAATGTGCCGAACCTAGAAA 213

Db 68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu 87

QY 214 GCCACCGAGTGGGTGAGGAATCCGATATTATTAAAGACAANTATCGTCCATATAAA 273

Db 88 GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys 107

QY 274 TTTAAAGTGACCTTTATTATAAAATATCATTCAGACGACGACATCGACGGGACGACA 333

Db 108 PheLysAlaAsnIleTyrTyrLysAsnIleIleIleThrThrValTrpSerGlySerThr 127

QY 334 TATACACAGATCAACTACGATATACAGATAGGACGCCGTTTCCATTGAGAGATCAGC 393

Db 128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147

QY 394 GATCTAAATCGACGGCAAGGAGATGCTCATCTAAAGCAAGATACCTTAGAACAACATGTA 453

Db 148 GluLeuIleAspArgArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr 167

QY 454 TATGTTGAAGCGTTTGACAGGATCGGGAGAAAAACAAGTACTTCTAAAACCATCAAAA 513

Db 168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187

QY 514 TTCACACGCGCGAATCTAGGCGATGGCACAGCTAATGAGACGTATACCGTGTGGGA 573

Db 188 PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly 207

QY 574 TCACCATGGATATATCGAACGGGAACCTCCGTCATTTAGTAGAGAAATGGATGCC 633

Db 208 GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla 227

QY 634 CGCTCTGTGTTTCCGTATTCATATTTCGAATGCCAATGGCGACATCGCGAACATATCT 693

Db 228 ArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspIleIleHisMetSer 247

QY 694 CCATTTTATGTCTATCCCCACAGAGCGTCCCGCAGAACCCATGGGATATCCCGAGAT 753

Db 248 ProphePheGlyLeuArg--AspGlyAlaHisThrGluTyrIleSerTyrSerThrAsp 266

QY 754 AATTTCAAACAACATAGATAGTATTTCATATGGATTGGACAGCGTCGAAAGACGAA 813

Db 267 ArgPheGlnGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGluGly 286

QY 814 CTTCCAGTCAAGCGTAACTTCTCATACATCACACTTCACAGTTGGTGGGACTGGGCT 873

Db 287 AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaIleTrpAsnTrpVal 306

QY 874 CCAAAAACACTACTCGTGTATGTTCAATGACTAAGTGGAAAGAGGTGACTGAATGTCCGT 933

Db 307 ProLysIleArgGluValCysThrLeuAlaLysTrpArgGluIleAspGluIleIleArg 326

QY 934 GCACAGTTAATGGCAATACAGATTTATGGCCGCTGAACCTTCGGCAACGTTTATCAGT 993

Db 327 AspGluTyrLysGlySerTyrArgPheThrAlaLysSerIleSerIleAlaThrPheIleSer 346

QY 994 AATACGACTGAGTTTGATCCAAATCGCATCATATTAGGACAANTGTATTAACCGGAGGCA 1053

Db	707	GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal	726
QY	2131	TTAGAGCAGCGCGAATGGAGCTTAACAGATCAATCAATGCTAAGCGATGACATTAGCT	2190
Db	727	LeuGlyAlaAlaA-----AsnAlaValIleAlaThrValSer	738
QY	2191	TCACTACAAGACTATGCCAAG-----ATTGAA	2217
Db	739	GlyValSerSerPheLeuAsnAsnPropheGlyAlaLeuAlaValGlyLeuIleLeu	758
QY	2218	GCTAGTTTATCATCTGCTTATAGTAGAAGCTGAACAGTTAACTTAACCTTAATGCAACA	2277
Db	759	AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro	778
QY	2278	TTAGAACACTAAATAATGGCTAAACT---AATTAGAAATCAGCCATCAACCAAGCTAAT	2334
Db	779	MetLysAlaLeuTyrProValThrThrArgAsnLeuLysGluSerValLysAsnGlyAsn	798
QY	2335	ACGGATAAACAGACTTTTGATAATGAACACCCAAATTTAGTTGAA-----	2379
Db	799	SerGlyAsnAsnSerAspGlyGluGluAsnAspAsnIleAspGluGluLysLeuGln	818
QY	2380	-----GCATACAAGCACTAAACCCACTTTAGAACACAGT-----	2415
Db	819	GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnHis	838
QY	2416	-----GCTACTAACCTTGAAGTTGTGCA	2439
Db	839	LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer	858
QY	2440	-----TCNACTGCTTATAATCAAAATTCGCAAT	2466
Db	859	LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn	871
RESULT 13			
ID	Q84735	PRELIMINARY;	PRT; 881 AA.
AC	Q84735;		
DT	01-NOV-1996	(TRENBLrel. 01, Created)	
DT	01-NOV-1996	(TRENBLrel. 01, Last sequence update)	
DT	01-DEC-2001	(TRENBLrel. 19, Last annotation update)	
DE	Glycoprotein B equivalent.		
CN	GB		
OS	Phocine herpesvirus type 1.		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphaherpesvirinae.		
OX	NCBI_TaxID=45104;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PB 84;		
RA	Harder T.C., Osterhaus A.D.M.E.;		
RT	"Identification, sequence analysis and immunogenicity of a marine		
RT	mammal herpesvirus (phocid herpesvirus-1) glycoprotein B equivalent		
RL	gene expressed in insect cells."		
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; Z68147; CRA92272.1; --		
DR	Interpro; IPR000234; Glycoprot_B.		
DR	Pfam; PF00606; Glycoprotein_B; 1.		
DR	ProDom; PD000693; Glycoprot_B; 1.		
SQ	SEQUENCE	881 AA; 99574 MW; FCB5B68EB35CC162 CRC64;	
Alignment Scores:			
Pred. No.:	2,46e-99	Length:	881
Score:	1930.00	Matches:	382
Percent Similarity:	63.38%	Conservative:	151
Best Local Similarity:	45.42%	Mismatches:	238
Query Match:	34.07%	Indels:	70
DB:	12	Gaps:	13
US-09-147-052-3 (1-3261) x Q84735 (1-881)			
QY	58	AACATCATCTCCGACTACCCAA---AATGTGACATCAACAGAAAGTTGTTCCGAGCGTCCAG	114

Db	38	AsnSerSerSerAsnThrGluLeuAsnAspMetArgGluIleLeuGlyGluSerGln	57
QY	115	TTG---TCTGAGGAAGTCTTACGTTTATCTTGTCCCCACACGAGTGGTTCACACGGT	171
Db	58	IleGluSerAspThrAlaThrPhePheMetCysProProSerGlySerThrLeu	77
QY	172	ATCCGCTAGAACCCGCCGGAATGTGCCGAACCTAGAAAAGCCACCGAGTGGGGTGA	231
Db	78	ValArgLeuGluProProArgAlaCysProAsnTyrLysLeuGlyLysAsnPheThrGlu	97
QY	232	GGAATCGGATATTTAAAGAGAATATCAGTCCATATAATTAAGTACCGCTTAT	291
Db	98	GlyIleAlaValIlePheLysGluAsnIleSerProTyrLysPheLysAlaAsnIleTyr	117
QY	292	TATAAATATCATTCAGACGACGACATGGAGGGGACGACATATAGACAGATCACTAAT	351
Db	118	TyrLysAsnIleIleIleThrThrValTrpSerGlySerSerTyrAlaValValThrAsn	137
QY	352	CGATATACAGATAGGACCCCGTTCCATTTCAAGAGATCAGGATCTTAATCGAGGCA	411
Db	138	MethIleThrAspArgValProIleLysValGlnGluIleThrGluLeuIleAspArg	157
QY	412	GGAAGATGCTCATCTAAAGCAAGATACCTTAGAACAAATGTATGTTGAAGCGTTTGC	471
Db	158	GlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyrGluPheThrAlaPheAsp	177
QY	472	AGGGATGGGGAGAAAACAATGCTCTTAAACCATCAAAATTCACACACCCGCAATCT	531
Db	178	LysAspGluAspProArgGluMetHisLeuLysProSerLysPheAsnThrProGlySer	197
QY	532	AGGCGATGGCACGACTAATGACAGCTATACCGTGTGGGATCACCATGATATATCGA	591
Db	198	ArgGlyTrpHisThrThrAsnAspThrTyrThrLysIleGlySerProGlyPheTyrArg	217
QY	592	ACGGAACTCCGTCATTTGATAGTAGAGAAATGGATGCCCTCTGTCTTCCGAT	651
Db	218	ThrGlyThrSerValAsnValGluValAlaAspAlaArgSerValTyrProTyr	237
QY	652	TCATATTTTGCATGGCAATGGGACATCCGCAACATATCTCCATTTTATGGTATACC	711
Db	238	AspSerPheGlyIleSerThrGlyAspIleHisMetSerProPhePheGlyLeuArg	257
QY	712	CCACGAGGCTGCCGCGAGAACCCATGGGATATCCCGAGGATATTTCAACACATAGAT	771
Db	258	---AspGlyAlaHisThrGluHisThrSerTyrSerAsnAspArgPheGlnIleGlu	276
QY	772	AGCTATTTTCAATGGATTTGGCAAGGTCGAAAAGCAAGCCCTCCAGTCAACGCTAAC	831
Db	277	GlyTyrTyrProIleAspLeuAspThrArgLeuGlnValGlyGlyProValSerArgAsn	296
QY	832	TTTCTCATCATCACACTTCACAGTTGGTGGGACTGGGCTCCCAAAACTACTCGTGA	891
Db	297	PheLeuThrThrGlnHisValThrValAlaTrpAsnTrpValProLysIleArgGluVal	316
QY	892	TGTTCAATGACTAAGTGGAAAGAGTGAAGTCAATGTTCCGTCGACAGTTAATGGGGA	951
Db	317	CysThrLeuAlaLysTrpArgGluIleAspGluIleIleArgAspGluTyrLysGlySer	336
QY	952	TACAGATTTATGGCCGCTGGAATTTTCGGCAACGTTTATCAGTATACGACGTAGTTGAT	1011
Db	337	TyrArgPheThrAlaLysSerIleSerAlaThrPheIleSerAspAlaThrGlnPheAsp	356
QY	1012	CCAAATCGCATCATATTAGGACAATGTATTAAACGAGGAGGAGGAGGAGGAGGAG	1071
Db	357	IleAsnArgValLysLeuSerAspCysAlaLysArgGluAlaThrGluAlaIleAspLys	376
QY	1072	ATATTTAGGACAAAATATATGACAGCTACCTCAAGTTGGAGTGCATGTAATATTCGTG	1131
Db	377	IleTyrLysAsnLysTyrAsnLysThrHisIleGlnThrGlyGluLeuGluThrTyrLeu	396
QY	1132	GCTCTCGGGGATTTATGTAGCATATCAGCTGTTCTTATCCAAATCCCTGGCTCATATG	1191
Db	397	AlaArgGlyGlyPheIleIleAlaPheArgPrometIleSerAsnGluLeuAlaLysLeu	416

```

QY 1192 TACCTCAGAGATTGATGAGACACACAGCAGCATGCTGCTGACCTGGTAAACAAT 1251
Db 1192 TACCTCAGAGATTGATGAGACACACAGCAGCATGCTGCTGACCTGGTAAACAAT 1251
QY 417 TyriLeuLeuLeuAlaArgSerGluArgIleValAspLeuAsnAlaLeuLeuAsnPro 436
Db 417 TyriLeuLeuLeuAlaArgSerGluArgIleValAspLeuAsnAlaLeuLeuAsnPro 436
QY 1252 AAGCATGCAATT-----TATAAGAAAATGCTACCTCATTTGTCAGGATTGCGCGAGAT 1305
Db 1252 AAGCATGCAATT-----TATAAGAAAATGCTACCTCATTTGTCAGGATTGCGCGAGAT 1305
QY 437 SerHisSerValGlyArgGlyLysArgSerIleGluThrGluThrLeuGlyArgSer 456
Db 437 SerHisSerValGlyArgGlyLysArgSerIleGluThrGluThrLeuGlyArgSer 456
QY 1306 ATTCGAATGCCCAATAGAAAATAACATTAGACGACACACAGCTATTAATCGACA 1365
Db 1306 ATTCGAATGCCCAATAGAAAATAACATTAGACGACACACAGCTATTAATCGACA 1365
QY 457 LysArgGlyValAspGlyValGlnAsnValAsnAlaThrLeuIleLysThrThr 476
Db 457 LysArgGlyValAspGlyValGlnAsnValAsnAlaThrLeuIleLysThrThr 476
QY 1366 TCGTCTGTTCAATTGCGCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATAT 1425
Db 1366 TCGTCTGTTCAATTGCGCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATAT 1425
QY 477 SerSerIleHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsn 496
Db 477 SerSerIleHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsn 496
QY 1426 GATATGTTTATAGGATTGCCAGCTTGGTGGAAATTCAGAAATAGAGAACTTGTTT 1485
Db 1426 GATATGTTTATAGGATTGCCAGCTTGGTGGAAATTCAGAAATAGAGAACTTGTTT 1485
QY 497 GluMetLeuSerArgIleAlaThrAlaTyrCysAsnLeuGlnAsnLysGluArgThrLeu 516
Db 497 GluMetLeuSerArgIleAlaThrAlaTyrCysAsnLeuGlnAsnLysGluArgThrLeu 516
QY 1486 TGGCAGCAAGGGATAAGATTATCTAGCGGTACAGCGATGCAACATTAGGAAGAGA 1545
Db 1486 TGGCAGCAAGGGATAAGATTATCTAGCGGTACAGCGATGCAACATTAGGAAGAGA 1545
QY 517 TrpAsnGluValMetLysLeuAsnProThrSerIleThrSerThrIleMetAspGlnLys 536
Db 517 TrpAsnGluValMetLysLeuAsnProThrSerIleThrSerThrIleMetAspGlnLys 536
QY 1546 GTGCTGCAAGATGTTGGGGATGTCGCTGCTGATCGAGCTGCACCTGCTATAGATCG 1605
Db 1546 GTGCTGCAAGATGTTGGGGATGTCGCTGCTGATCGAGCTGCACCTGCTATAGATCG 1605
QY 537 ValSerAlaArgLeuLeuGlyAspValIleAlaValThrGlnCysValAsnIleSerGly 556
Db 537 ValSerAlaArgLeuLeuGlyAspValIleAlaValThrGlnCysValAsnIleSerGly 556
QY 1606 GAATCGCTCACTTTGCAAAATTTCTATCGAGTTTATCACATCCACTATATATATAGC 1665
Db 1606 GAATCGCTCACTTTGCAAAATTTCTATCGAGTTTATCACATCCACTATATATATAGC 1665
QY 557 SerAsnValPheIleGlnAsnSerMetArgValThrGlySerThrThrThrCysTyrSer 576
Db 557 SerAsnValPheIleGlnAsnSerMetArgValThrGlySerThrThrThrCysTyrSer 576
QY 1666 CGACCATGCTTCTATTTTCATATGAGAAACCAACAGGAAC-----ATACAGGACAACATC 1722
Db 1666 CGACCATGCTTCTATTTTCATATGAGAAACCAACAGGAAC-----ATACAGGACAACATC 1722
QY 577 ArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAspTyrIleGluGlnLeu 596
Db 577 ArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAspTyrIleGluGlnLeu 596
QY 1723 GTGTAACAACAGAGTTGCTCCACGCTAGAGCTGCTAGAGCTGCTGCGCTATCAT 1782
Db 1723 GTGTAACAACAGAGTTGCTCCACGCTAGAGCTGCTAGAGCTGCTGCGCTATCAT 1782
QY 597 GlyGluAsnAsnGluLeuLeuValAspArgLysLeuIleGluProCysThrAlaAsnAsn 616
Db 597 GlyGluAsnAsnGluLeuLeuValAspArgLysLeuIleGluProCysThrAlaAsnAsn 616
QY 1783 CGTAGATATTTCTGTTTGGATCCGGTTATGCTTTATTTGAAACTATATTTGTTAAG 1842
Db 1783 CGTAGATATTTCTGTTTGGATCCGGTTATGCTTTATTTGAAACTATATTTGTTAAG 1842
QY 617 LysArgTyrPheLysPheGlyValAspTyrValTyrPheGluAsnTyrValTyrIleArg 636
Db 617 LysArgTyrPheLysPheGlyValAspTyrValTyrPheGluAsnTyrValTyrIleArg 636
QY 1843 ATGTAGAGCTGCCCATATACAGATTGCTAGACATTTGTCGAGCTTAATCTAACCCG 1902
Db 1843 ATGTAGAGCTGCCCATATACAGATTGCTAGACATTTGTCGAGCTTAATCTAACCCG 1902
QY 637 LysValProLeuAsnGluLeuGluMetIleSerThrTyrValAspLeuAsnIleThrLeu 656
Db 637 LysValProLeuAsnGluLeuGluMetIleSerThrTyrValAspLeuAsnIleThrLeu 656
QY 1903 CTAGAAGATCGGAAATTTGCTTTATCCGTTTACACAAAAGAGAGTTCGCTGATGTT 1962
Db 1903 CTAGAAGATCGGAAATTTGCTTTATCCGTTTACACAAAAGAGAGTTCGCTGATGTT 1962
QY 657 LeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThr 676
Db 657 LeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThr 676
QY 1963 GGTGATTGATATGACAGAACTGCTGCCCATCACTCACTCACTCACTCACTCACT 2022
Db 1963 GGTGATTGATATGACAGAACTGCTGCCCATCACTCACTCACTCACTCACTCACT 2022
QY 677 GlyLeuLeuAspTyrSerGluTyrIleGlnArgArgAsnGlnLeuHisAlaLeuLysPheTyr 696
Db 677 GlyLeuLeuAspTyrSerGluTyrIleGlnArgArgAsnGlnLeuHisAlaLeuLysPheTyr 696
QY 2023 GACATAAACAAGTAATAGAACTGGATACAAAT----- 2055
Db 2023 GACATAAACAAGTAATAGAACTGGATACAAAT----- 2055
QY 697 AspIleAspSerValValLysValAspAsnAsnLeuIleIleMetArgGlyMetLeuThr 716
Db 697 AspIleAspSerValValLysValAspAsnAsnLeuIleIleMetArgGlyMetLeuThr 716
QY 2056 ----TACGGGGCTCGAGAAATTCGGCTGATGCTATCTATTACTATAAAGATGCAACCCA 2112
Db 2056 ----TACGGGGCTCGAGAAATTCGGCTGATGCTATCTATTACTATAAAGATGCAACCCA 2112
QY 717 PhePheGlnGlyLeuGlyAspValGlyAla----- 726
Db 717 PhePheGlnGlyLeuGlyAspValGlyAla----- 726
QY 2113 ATATATGCCCAACCCATAGAACAGCGGGAATGGAGTTAACAGATCTAATCAATGCT 2172
Db 2113 ATATATGCCCAACCCATAGAACAGCGGGAATGGAGTTAACAGATCTAATCAATGCT 2172
QY 727 GlyPheGlyLysValValLeuGlyAlaAla-----AsnAla 738
Db 727 GlyPheGlyLysValValLeuGlyAlaAla-----AsnAla 738
QY 2173 AAAGCGATGATAGTTCATACAGACATGATGCCAAG----- 2211
Db 2173 AAAGCGATGATAGTTCATACAGACATGATGCCAAG----- 2211
QY 739 ValIleSerThrValSerGlyIleSerSerPheLeuAsnAsnProPheGlyAlaLeuAla 758
Db 739 ValIleSerThrValSerGlyIleSerSerPheLeuAsnAsnProPheGlyAlaLeuAla 758

```

```

QY 2212 -----ATTGAAGCTAGTTTATCATCTGCTTTATAGTGAAGCTGAACAGTTAAC 2259
Db 759 ValGlyLeuLeuIleLeuAlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSer 778
QY 2260 AATAACCTTAATCAACATTAGAACAACTA-----AAATGGCTAAACT 2304
Db 779 LysLeuLysSerAsnProMetLysAlaLeuTyrProValThrThrArgAsnLeuLysGlu 798
QY 2305 AATTGATGATCAAGCAACCAAGCTAATACGATATAAAGCACTTTTGTATATGAA--- 2361
Db 799 SerSerLysGluLysIleGlyAspGlyAspGluAspGlyAspGluPheAspGluAspLys 818
QY 2362 -----CACCCAAATTTAGTTGAAGCA----- 2382
Db 819 LeuSerGlnAlaLysGluMetIleLysTyrMetThrLeuIleSerAlaMetGluLysGln 838
QY 2383 ---TACAAAGCACTAAAA-----ACCACCTTAGAACAAAGCTGCTACTAAC 2424
Db 839 GluHisLysAlaMetLysLysAsnSerGlyProAlaIleLeuAlaAsnArgValAlaAsn 858
QY 2425 CTT 2427
Db 859 Leu 859

RESULT 14
Q90050 PRELIMINARY; PRT; 948 AA.
AC Q90050;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Herpes simplex virus glycoprotein B (gB) homolog.
OS Feline herpesvirus (feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93090104; PubMed=1333759;
RA Maeda K., Horimoto T., Norimine J., Kawaguchi Y., Tomonaga K.,
RA Nikiura M., Kai C., Takahashi E., Mikami T.;
RT "Identification and nucleotide sequence of a gene in feline
RT herpesvirus type 1 homologous to the herpes simplex virus gene
RT encoding the glycoprotein B";
RL Arch. Virol. 127:387-397(1992).
DR EMBL; S49775; AAB24381.2; -
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00606; Glycoprotein_B.
DR ProDom; PD000693; Glycoprotein_B; 1.
DR PROSITE; PS00018; EF_HAND; 1.
SQ SEQUENCE 948 AA; 106233 MW; 71C28FB0B091325 CRC64;

Alignment Scores:
Pred. No.: 3,398-98 Length: 948
Score: 1909.50 Matches: 391
Percent Similarity: 59.08% Conservative: 162
Best Local Similarity: 41.77% Mismatches: 256
Query Match: 33.71% Indels: 127
DB: Gaps: 14

US-09-147-052-3 (1-3261) x Q90050 (1-948)
QY 7 TATTTTACGGCAATTCATATTT----- 30
Db 22 TyrPheArgGlnArgCysPhePheSerLeuLeuGlyIleAlaThrGlySerArg 41
QY 31 -----TTCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTyrIle 61
QY 43 ATTTATAT----- 51
Db 43 ATTTATAT----- 51

```

Db	62	ValLeuPheLeuValGlyProArgProValGluGlyGlnSerThrSerGluGln	81
QY	52	-----GGTACGAATCATCTCCGAGTACCCAA	78
Db	82	ProArgArgThrValAlaThrProGluValGlyThrProProLysProThrThrAsp	101
QY	79	AATGTGACATCAAGAGAAGTTGTCGAGCGTCCAGTCTCTCAGGAACAG	129
Db	102	ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly	121
QY	130	---TCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTGCATCCGCTAGAACGG	186
Db	122	ProSerThrPheTyrMetCysProProSerGlySerThrValValargLeuGluPro	141
QY	187	CGCGAANAATGCCCAACTAGAAAGCCACCGAGTGGGTGAAGGAATCGCATATTA	246
Db	142	ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle	161
QY	247	TTTAAAGACAATACAGTCCATTAATTAAGTCACCGCTTTATTAATAAATATCAT	306
Db	162	PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle	181
QY	307	CAGACGACGACATGACGGGGGACATATAGACAGATCACTAATCGATATACAGATGG	366
Db	182	MetThrThrValTyrSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg	201
QY	367	ACGCCGTTTCATGAAGAGATCAGGATCAATCGAGCGCAAGGAAGATGCTCATCT	426
Db	202	ValProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSer	221
QY	427	AAAGCAAGATACCTTAGAACAACTATATGTTGAAGCGTTTCACAGGGATCGGGAGAA	486
Db	222	LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro	241
QY	487	AAACAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGCGATGGCACAG	546
Db	242	ArgGluLeuProLeuLysProSerLysPheAsnThrProGluSerArgGlyTyrPheIleThr	261
QY	547	ACTAATGACAGTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACCTCCGTC	606
Db	262	ThrAsnGluThrTyrThrLysIleGlyAlaAlaGlyPheHisSerGlyThrSerVal	281
QY	607	AATGTATATAGAGGAATGATGCCCTCTGTTGTTCCGTTATTCATATTTTGAATG	666
Db	282	AsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIle	301
QY	667	GCACATGGCGACATCCGACATATCTCCATTTTATGCTCTATCCCCACAGAGGCTGCC	726
Db	302	SerThrGlyAspValIleHisMetSerProPhePheGlyLeuArg---AspGlyAlaHis	320
QY	727	GCAGAACCCATGGGATATCCCGAGGATAATTCAAACAACTAGATAGCTATTTTCAATG	786
Db	321	ValGluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrTyrProIle	340
QY	787	GATTTGGACACGGCTCGAAACCAAGCCCTTCCAGTCAAGCTTAACCTTTCTCATCATCA	846
Db	341	AspLeuAspThrArgLeuGlnLeuGlyAlaProValSerArgAsnPheLeuGluThrPro	360
QY	847	CACCTTCACAGTTGGGTGGGCTCCAAAACCTACTCGTGTATGTTCAATGACTAAG	906
Db	361	HisValThrValAlaThrAsnThrPheProLysCysGlyArgValCysThrLeuAlaLys	380
QY	907	TGGAAGAGGTGACTGAATGTGCGTCAACAGTTAATGGGAGATACAGATTTATGGCC	966
Db	381	TrpArgGluIleAspGluMetLeuArgAspGluTyrGlnGlySerTyrArgPheThrVal	400
QY	967	CGTGAACCTTCGGCAAGCTTTATCAGTATACGACTGAGTTTCATCCAAATCGCATCAT	1026
Db	401	LysThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArg	420
QY	1027	TTAGGCAATGTATTAAACCGCAGCAAGCAAGCAGCATGTTGACAGATATTTAGGACAAA	1086
Db	421	LeuGlyAspCysAlaThrLysGluAlaAlaGluAlaIleAspArgIleTyrLysSerLys	440
QY	1087	TATAATGACAGTCACGTCACGTTGTCATGATGACATGACAAATATTTCTGCTCTCGGGGATTT	1146
Db	441	TyrSerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaargGlyGlyPhe	460
QY	1147	ATTGTAGCATATACGCTGCTTCATCCAAATCCCTCGCTCATATGTCACCTCAGAGATTC	1206
Db	461	LeuIleAlaPheArgPrometIleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeu	480
QY	1207	ATGAGAGACAACAGGACCGATGAGATGCTCGACCTGTAACAAATAAGCATGCAATTTAT	1266
Db	481	AlaArgSerAsnArgThrValAspLeuSerAlaLeuLeuAsn-----	494
QY	1267	AGAAAATATGCTACCTTCATTCACCATTCGCGGCGAGATATTCGAATGCAACCAATAGA	1326
Db	495	---ProSerGlyGluThrValGlnArgThrArgArgSerValProSerAsnGlnHisHis	513
QY	1327	AAAAATA-----ACATTACACACACACAGCTATT	1356
Db	514	ArgSerArgArgSerThrIleGluGlyGlyIleGluThrValAsnAsnAlaSerLeuLeu	533
QY	1357	AAATCGACATCGTCTCTCAATTCGCGATGCTCCAAATTTCTTTATGATCATATACAAACC	1416
Db	534	LysThrThrSerSerValGluPheAlaMetLeuGlnPheAlaTyrAspTyrIleGlnAla	553
QY	1417	CATATTATGATATGTTTAGTAGGATGCCACAGCTTGGTGGCAATTCGAGAATAGAGAA	1476
Db	554	HisValAsnGluMetLeuSerArgIleAlaThrAlaTyrCysThrLeuGlnAsnArgGlu	573
QY	1477	CTTGTATTATGCCACCAAGGATAAAGATTATCTAGCCCTACACGGATGCAACATTA	1536
Db	574	HisValLeuTyrPheThrGluThrLeuLysLeuAsnProGlyGlyValValSerMetAlaLeu	593
QY	1537	GGAAGGAGATGCTGCTCAAGATGTTGGGGATGTCGCTGTATCGAGCTCACTGCT	1596
Db	594	GluArgArgValSerAlaargLeuLeuGlyAspAlaValAlaValThrGlnCysValAsn	613
QY	1597	ATAGATCGCGAATCCCTCACTTTCGAAAATTCATCGGATTTATCATCCATCAATACA	1656
Db	614	IleSerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrThr	633
QY	1657	TGTTATAGCGCACATTTGTTCTATTTCATATGGAAGAAACCAAGAAACATACAGGA	1716
Db	634	CystSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluTyrIleGluGly	653
QY	1717	CAACTCGTGAACCAACAGGATGCTTCCAAACCTAGAGGCTAGAGCATCTCGGCT	1776
Db	654	GlnLeuGlyGluAsnAsnAspLeuLeuValGluArgLysLeuIleGluProCysThrVal	673
QY	1777	AATCATCGTAGATATTTCTGTTTGGATCCGGTTATGCTTTATTTGAAAACATAAATTT	1836
Db	674	AsnAsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyr	693
QY	1837	GTAAAGATGATAGACCTGCCGATATACAGATTCATAGACATTTGTCGAGCTTAATCTA	1896
Db	694	ValArgLysValProLeuSerGluIleGluLeuIleSerAlaTyrValAspLeuAsnLeu	713
QY	1897	ACCTCTAGAAGATCGGAAATTTTCCTTTATTCGTTTACACAAAAGAGAGTTCGCT	1956
Db	714	ThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGlu	733
QY	1957	GATGTTGTTGTTATGATTATGAGAAGTAGCTCGCCGCAATCACTACATCAATGCTAAA	2016
Db	734	AspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuLys	753
QY	2017	TTTTATGACATAACAAAGTAATAGAGTGGATACAAAT-----	2055
Db	754	PheTyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMet	773
QY	2056	-----TACGCGGCTGCGAGGAATTCGGCTGT-----	2082
Db	774	AlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGly	793

```

QY 2083 -----ATGCTATTACTAAAAAGATGCAAAACCCAAATTAATGCGCA 2124
Db 794 AlaAlaSerAlaValIleSerThrValSerGlyValSerSerLeuAsnAsnProPhe 813
QY 2125 ACCCAATTAGACGACGCGAATGAGTAAACAGATCTAATCAATGCT- 2172
Db 814 GlyAlaLeuAlaValGlyLeuLeuLeuAlaGlyIleValAlaAlaPheLeuAlaTyr 833
QY 2173 -----AAAGCGATG-----ACATTAGCT 2190
Db 834 ArgTyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArg 853
QY 2191 TCACATACAGACTAGTCCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 2250
Db 854 AsnLeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyGlyAspSerAspProGly 873
QY 2251 ACAGTTAAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAA- 2307
Db 874 ValAspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSer 893
QY 2308 TTAGATCAGCATCAACCAAGCTAATACGATTAACAGACTTTTGTATTAATGAACACCCA 2367
Db 894 LeuValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyPro 912
QY 2368 AATTAGTTGAAGCATAC-----AAAGCACTAAAAACC-----ACTTTAGAACAA 2412
Db 913 AlaIleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgGlyProLysTyrGln 932
QY 2413 CGTCTACTAACCTTGAAGGTTTGTCACTCAACTGCTTATAATCAAAAT 2460
Db 933 ArgLeuAsnAsnLeuAspSerGlyAspAspThrGluThrAsnLeuVal 948

RESULT 15
Q9QAP5 PRELIMINARY; PRT; 919 AA.
AC Q9QAP5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DE 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glycoprotein B.
OS Caprine herpesvirus 1 (goat herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=39944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E/CH;
RX MEDLINE=9221732; PubMed=10203465;
RS Ros C., Belak S.;
RT "Studies of genetic relationships between bovine, caprine, cervine,
RT and rangeliferine alphaherpesviruses and improved molecular methods for
RT virus detection and identification."
RL J. Clin. Microbiol. 37:1247-1253(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E/CH;
RS Ros C., Belak S.;
RT "Characterization of the glycoprotein B gene from ruminant
RT alphaherpesviruses."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF078728; AAD46114.2; -
DR InterPro; IPR00234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein.B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.
SQ SEQUENCE 919 AA; 100172 MW; A7DDCF03D13B4B66 CRC64;

```

Alignment Scores:
Pred. No.: 2,13e-92 Length: 919
Score: 1805.00 Matches: 354
Percent Similarity: 66.08% Conservative: 98
Best Local Similarity: 51.75% Mismatches: 208
Query Match: 31.86% Indels: 24
DB: 12 Gaps: 7

```

US-09-147-052-3 (1-3261) x Q9QAP5 (1-919)
QY 67 CCGAGTACCCAAATGTGACATCA-----AGAA 93
Db 83 ProSerAlaAspGlyAlaThrAlaAlaAlaGlyAsnGluThrAlaAlaAspValArg 102
QY 94 GAAAGTGTTCAGCGCTGAGTGTCTGAGGAAGAGTCTACGTTTATCTTTTGTCTCCCA 153
Db 103 AlaValLeuArgGluAlaGlnAlaAlaGlyGluHisSerGlnPhePheValCysProPro 122
QY 154 CCAGTGGTTCACCGCTGATCCGCTAGAACCCGCGGAAATATTCGCCGAACCTAGAAAA 213
Db 123 ProSerGlyAlaThrValValArgLeuAlaProGlyArgProCysProGluTyrGluLeu 142
QY 214 GCCACCGAGTGGGTGAAGGAATCCGATATATTTAAAGAGAATATCAGTCCATATAA 273
Db 143 GlyArgAsnTyrThrGluGlyIleGlyIleTyrLysGluAsnIleAlaProTyrLys 162
QY 274 TTTAAAGTACGCTTTTATATAAAATATCATTCAGACGACGACATGACGCGGAGCACA 333
Db 163 PheLysAlaHisIleTyrTyrLysAsnValIleValThrThrThrThrSerGlySerThr 182
QY 334 TATAGACAGATCACTAATTCGATATACAGATAGGACGCCGCTTCCATTGAAGACATCAG 393
Db 183 TyrAlaAlaIleThrAsnGlnHisThrAspArgValProValGlyValSerGluIleThr 202
QY 394 GATCTAATCGACGCGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTTAGAACATGTA 453
Db 203 AspValValAspLysLysTyrPargCysLeuSerLysAlaGluTyrLeuArgSerGlyArg 222
QY 454 TATGTTGAAGCGTTTGACAGGATCGCGGAGAAAAACAAGTACTTCTAAAACCTCAAAA 513
Db 223 LysValValAlaPheAspArgAspAlaAspProThrGluAlaProLeuLysProSerArg 242
QY 514 TTCACACGCGCGAATCTAGGCGATGCGACAGCACTAATGAGACGTATACCGTGTGGGA 573
Db 243 LeuAsnAlaProGlyAlaArgGlyTyrHisThrThrAspGluValHisThrValValGly 262
QY 574 TCACCATGGATATATCGAACGGAACCTCGCTCAATTCATAGTAGAGGAATGGATGCC 633
Db 263 SerValGlyLeuTyrArgThrGlyThrSerValAsnCysIleValGluGluValGluAla 282
QY 634 CGCTCTGTGTTCCGTATTCATATTTGCAATGGCCAAATGGCGACATCGCGAATATCT 693
Db 283 ArgSerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleIleTyrMetSer 302
QY 694 CCATTTTATGTTCTATCCCGACAGAGGCTGCCCGACAGAACCCATGGGATATCCCGAGAT 753
Db 303 ProPheTyrGlyLeuArg---GluGlyAlaHisArgGluHisThrSerTyrSerProAsp 321
QY 754 AATTTCAAACAACCTAGATAGCTATTTTTCATGATTTGGACAAGCTCGCAAGCAAGC 813
Db 322 ArgPheGlnGlnIleGluGlyTyrTyrSerArgAspLeuThrSerGlyArgArgAlaGly 341
QY 814 CTTCAGTCAAGCGTAACCTTCTCATCATCATCATCATCTTCACAGTTGGGTGGGATGGCT 873
Db 342 AspProValSerArgAsnPheLeuArgThrGlnHisValThrValAlaTyrAspTyrVal 361
QY 874 CCAAAAACTACTCGTGTATGTTCAATGACTAAGTGGAAAGAGGTGACTGAATGTTCCGT 933
Db 362 ProLysArgLysAsnValCysSerLeuAlaArgTyrPargAlaAlaGluGluMetLeuArg 381
QY 934 GCAACAGTTAATGGGAGATACAGATTTATGCGCGCTGCAACTTTCGGCAACGTTTATCAGT 993
Db 382 AspGluSerArgGlyAsnTyrArgPheThrAlaArgAlaLeuSerAlaThrPheValSer 401
QY 994 AATACGACTGAGTTTGTATCCAAATCGCATCATATTTAGGACAATGTATTAACCGAGGCA 1053
Db 402 AspThrHisAlaPheSerLeuGlnHisValProLeuSerAspCysValLeuGluAspAla 421
QY 1054 GAAGCAGCAATCGACGAGATATTTAGGACAAAATATATATACACAGTACAGTTGGA 1113

```

Db	422	GluAlaGluValGluArgValTyrArgGluArgTyrAsnGlyThrHisValLeuAlaGly	441
Qy	1114	CATGTACAATATTCTTGGCTCTCGGGGGATTATTGTAGCATATCAGCTGTCTATCC	1173
Db	442	LysThrAspThrTyrLeuAlaArgGlyGlyPheValAlaPheArgProLeuLeuSer	461
Qy	1174	AAATCCCTGGCTCATATGTACCTCAGAGAATTGATGAGAGACAACAGGACCGATGAGATG	1233
Db	462	AsnGluLeuAlaLysLeuTyrLeuGlnGluLeuAlaArgSerAsnArgThr	478
Qy	1234	CTCAGCTGCTAAACAATAAGCATGCAATTTATTAAGAAAAATCTACTCTATTGTCCAGG	1293
Db	479	LeuAspGlyValLeuGlyProArg-----LeuProAlaGlyAlaAlaProAlaGlyAla	496
Qy	1294	TTGGGGCGGAGATATTCGAATGTGACCAAAATAGAAAAATACATTAGACGACACACAGCT	1353
Db	497	LeuArgArgAlaArgArgAlaAlaProGlyGly-----GlyGlyAlaGlyArg	513
Qy	1354	ATTAAATCGACATCGCTGTTCAAATTCGCCCATGCTCCAATTTCTTTATGATCATATACAA	1413
Db	514	ValThrValSerSerAlaGluPheAlaAlaLeuGlnPheThrTyrAspHisIleGln	533
Qy	1414	ACCATATTATGATATTTTAGTAGGATTGCCACAGCTTTGGTCGGAATTCAGAGATAGA	1473
Db	534	AspHisValAsnAlaMetPheSerArgLeuAlaThrSerTrpCysLeuLeuGlnAsnLys	553
Qy	1474	GAACCTGTTTATGGCAGGAAGGATTAAGATTAACTCTAGCGCTCAGCGGATGCAACA	1533
Db	554	GluArgAlaLeuTrpAlaGluAlaAlaLysLeuAsnProSerAlaAlaLaserAlaAla	573
Qy	1534	TTAGGAGGAGAGTGGCTGCAAAAGATGTTGGGGATGTCGCTGCTATCGAGCTGCACCT	1593
Db	574	LeuAsnArgArgAlaAlaAlaArgMetLeuGlyAspAlaMetAlaValThrTyrCysArg	593
Qy	1594	GCTATAGATCGGAATCCGCTACTTTTGCAAAATTCATTCGGAGTTATTCACATCCACTAAT	1653
Db	594	GluLeuGlyAlaGlyArgValPheIleGluAsnSerMetArg---MetProGlyGlyAla	612
Qy	1654	ACATGTTATAGCCGACCATTTGGTCTATTTTCATATGGAGAAAACCAAGAAACATACAG	1713
Db	613	AlaCysTyrSerArgProLeuValSerPheAlaTyrGlyAsnGluSerGluProLeuGlu	632
Qy	1714	GGCAACTCGGTGAAACAACAGATGTTCTTCCACAGCTAGAGGCTGTAGAGCCATGCTCG	1773
Db	633	GlyGlnLeuGlyGluAspAsnGluLeuLeuLeuSerArgAspLeuValGluProCysAla	652
Qy	1774	GCTAATCATCGTAGATATTTCTGTTTGGATCGGTTATGCTTTATTTGAAACTATTAAT	1833
Db	653	AlaAsnHisArgArgTyrPheArgPheGlyAlaAspTyrValTyrTyrGluAsnTyrAla	672
Qy	1834	TTTGTTAAGATGTAGACGCTGCCGATATACAGATTGCTAGCACATTGTTCGAGCTTAAT	1893
Db	673	TyrValArgArgValProLeuThrGluIleGluThrIleSerThrPheValAspLeuAsn	692
Qy	1894	CTAACCTGCTAGAGATCGGGAAATTTTGCCCTTTATTCGTTTACACAAAGAAGATTG	1953
Db	693	LeuThrValLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeu	712
Qy	1954	CGTGATCTTGGTGTATGGATTATGCAGAGTAGCTCGCGCAATCACTACATGAACCTT	2013
Db	713	AlaAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisGluLeu	732
Qy	2014	AAATTTTATGACATAAACAAGTAATAAGAGTGGATACAAATTTACGG-----GGG	2064
Db	733	LysPheThrAspIleAspArgValValLysLeuAspSerAsnMetValIleMetArgGly	752
Qy	2065	CTCAGGAATTC	2076
Db	753	LeuAlaSerPhe	756

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:33:17 ; Search time 52.5576 Seconds
(without alignments)
5835.661 Million cell updates/sec

Title: US-09-147-052-3
Perfect score: 5665
Sequence: 1 atgcactatttagcgagaa.....gaccgggtacattttataa 3261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09147052/runat_08102003_154340_29835/app_query.fasta_1.5980
-DB=SwissProt.41 -QFMT=fastan -SUFFIX=rs -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052.ecgn_1.1_140 -runat_08102003_154340_29835 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3601	63.6	865	1 VGLB_HSVMD	P18538 marek's dis
2	1785.5	31.5	980	1 VGLB_HSVBA	P18551 equine herp
3	1779.5	31.4	980	1 VGLB_HSVBA	P28922 equine herp
4	1777.5	31.4	868	1 VGLB_VZVD	P09257 varicella-z
5	1775	31.3	913	1 VGLB_PRTVF	P08355 pseudorab
6	1761.5	31.1	979	1 VGLB_HSVEL	P25218 equine herp
7	1760.5	31.1	919	1 VGLB_HSVBA	P17472 equine herp
8	1741.5	30.7	980	1 VGLB_HSVEL	P18050 equine herp
9	1736	30.6	904	1 VGLB_HSV2H	P08666 herpes simp
10	1735	30.6	904	1 VGLB_HSV2H	P06763 herpes simp
11	1731.5	30.6	932	1 VGLB_HSVBC	P12640 bovine herp
12	1729.5	30.5	920	1 VGLB_HSVSM	Q04464 herpesviru
13	1714.5	30.3	885	1 VGLB_HSV2S	P24994 herpes simp
14	1710	30.2	917	1 VGLB_HSVB2	P12641 bovine herp
15	1705.5	30.1	903	1 VGLB_HSVIF	P06436 herpes simp
16	1705.5	30.1	904	1 VGLB_HSV11	P10211 herpes simp
17	1696.5	29.9	904	1 VGLB_HSV1P	P08665 herpes simp
18	1693.5	29.9	904	1 VGLB_HSV1K	P06437 herpes simp

19	1675.5	29.6	933	1 VGLB_HSV1	Q04463 herpesvirus
20	1604	28.3	928	1 VGLB_HSVBP	P17471 bovine herp
21	1337.5	23.6	883	1 VGLB_ILTVS	P27415 infectious
22	1325	23.4	883	1 VGLB_ILTVT	P24904 infectious
23	1324	23.4	873	1 VGLB_ILTV6	Q02409 infectious
24	778.5	13.7	907	1 VGLB_HCMVT	P13201 human cytom
25	758	13.4	830	1 VGLB_HSV6U	P28864 human herpe
26	756	13.3	906	1 VGLB_HCMVA	P06473 human cytom
27	750	13.2	830	1 VGLB_HSV6G	P36319 human herpe
28	741	13.1	822	1 VGLB_HSV7J	P52352 human herpe
29	739	13.0	830	1 VGLB_HSV62	P36320 human herpe
30	738	13.0	901	1 VGLB_GPCMV	Q69024 guinea pig
31	737	13.0	928	1 VGLB_MCMVS	P27171 murine cyto
32	717	12.7	854	1 VGLB_RHCM6	P89053 rhesus cyto
33	709.5	12.5	857	1 VGLB_EBV	P30188 Epstein-Bar
34	700.5	12.4	944	1 VGLB_HSVT2	Q9wr15 herpesvirus
35	695	12.3	808	1 VGLB_HSVSA	P24905 herpesvirus
36	185	3.3	1433	1 CAT8_YEAST	P39113 saccharomyc
37	185	3.3	2660	1 YEEJ_ECO57	Q8x8v7 escherichia
38	176	3.1	3712	1 LMA_DROME	Q00174 drosophila
39	172	3.0	1085	1 RBP2_PLAVB	P24339 schizosacch
40	170	3.0	1251	1 CUP7_SCHPO	Q00799 plasmodium
41	167.5	3.0	1398	1 MHPI_YEAST	P43638 saccharomyc
42	166.5	2.9	1093	1 TMFL_HUMAN	P82094 homo sapien
43	163.5	2.9	1276	1 BXD_CLOBO	P19321 clostridium
44	161	2.8	1630	1 MSP1_PLAFK	P04932 plasmodium
45	161	2.8	1639	1 MSP1_PLAFW	P04933 plasmodium

ALIGNMENTS

RESULT 1
ID VGLB_HSVMD STANDARD: PRT; 865 AA.
AC P18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_Taxid=33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293086; PubMed=2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;
RT "Nucleotide sequence and characterization of the Marek's disease
virus homologue of glycoprotein B of herpes simplex virus.";
RL J. Gen. Virol. 70:1789-1804(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13713; BAA02866.1; --
CC InterPro; IPR000234; Glycoprot.B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD000693; Glycoprot.B; 1.
CC Signal; Glycoprotein; Transmembrane.
CC SIGNAL 1 21 POTENTIAL.
CC CHAIN 22 865 GLYCOPROTEIN B.
CC DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 683 700 POTENTIAL.
CC TRANSMEM 709 729 POTENTIAL.
CC TRANSMEM 732 752 POTENTIAL.
CC DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).

FT	CARBOHYD	184	184	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	425	425	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	631	631	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	865 AA; 98091 MW; B30E93C1AC65C63 CRC64;			
Alignment Scores:					
Pred. No.:	2,22e-191	Length:	865		
Score:	3601.00	Matches:	719		
Percent Similarity:	87.81%	Conservative:	16		
Best Local Similarity:	85.90%	Mismatches:	60		
Query Match:	63.57%	Indels:	42		
DB:	1	Gaps:	8		
US-09-147-052-3 (1-3261) x VGLB_HSVMD (1-865)					
QY	1	ATGCACTATTATAGCGGAATTCGATATATTTTCCCTTATAGTTATTTCTATATGGTACGAAC	60		
DB	1	MetHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn	20		
QY	61	TCATCTCCAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT	120		
DB	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40		
QY	121	GAGGAAGAGTCTACGTTTATCTTTGTCCCCCACCAGTGGGTCAACCGTGATCCGCTCA	180		
DB	41	GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu	60		
QY	181	GAACGCCCGGAAATGTCGCCAACTAGAAAAGCCAGCGTGGGGTGAAGAAATCGCG	240		
DB	61	GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGlyIleAla	80		
QY	241	ATATATTTAAAGAGATATCATGTCATATAATTTAAAGTGACCGTTTATTATAAAAT	300		
DB	81	IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn	100		
QY	301	ATCATTCAGACGACATGAGCGGAGGAGCATATAGACATCATCATCATCATATACA	360		
DB	101	IleIleGlnThrThrTrpThrTrpGlyThrThrTrpArgGlnIleThrAsnArgTyrThr	120		
QY	361	GATAGGACCGCTTCCATTTGAAGAGATCACGGATCTAATCGACGGCAAGGAATGCG	420		
DB	121	AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys	140		
QY	421	TCATCTAAAGCAAGATACCTTGAACAAATGTATATGTTGAAGCGTTTGACAGGGATGCG	480		
DB	141	SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla	160		
QY	481	GGAGAAAACAGTACTTCTAAACCATCAAAATTCACACGCCCGGAATCTAGGGCATGG	540		
DB	161	GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180		
QY	541	CACACGACTAATCAGACGTATACCGTGGGGATCACCATGATATATATCAACCGGAAAC	600		
DB	181	HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr	200		
QY	601	TCGGTCAATGTATAGAGAAATGGATGCGCGCTCTGTGTTCCGCTATTCAATTTT	660		
DB	201	SerValAsnCysIleValGluMetAspAlaArgSerValPheProTyrSerTyrPhe	220		
QY	661	GCAATGGCCAAATGGCCACATCCGGAACATATCTCCATTTTATGTTGCTATCCCCCAGAG	720		
DB	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProGlu	240		
QY	721	GCTGCCGCAAGCCCATGGATATCCCGAGGATAATTTCAAAACATAGATAGCTATTTT	780		
DB	241	AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelysGlnLeuAspSerTyrPhe	260		
QY	781	TCAATGGATTGGACAGCGTCGAAAAGCAAGCCTTCCAGTCAAGCGTAACCTTTCATC	840		
DB	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280		

QY	841	ACATCACACTTTCACAGTTGGTGGGACTGGGCTCCAAACACTACTCGTGTATGTTCAATG	900		
DB	281	ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet	300		
QY	901	ACTAGTGGAAAGAGGTGACTGAATGTTGGTGCACACAGTTAATGGGAGATACAGATT	960		
DB	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320		
QY	961	ATGGCCCGTGAACCTTTCGGCAACGTTTATCAGTATACGACTGAGTTTGTATCAATCGC	1020		
DB	321	MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340		
QY	1021	ATCATATTAGGACAATGTATTAAACGCGAGCGAGCAAGCAACATCGACAGCATATTAGG	1080		
DB	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360		
QY	1081	ACAAAATATATGACAGTCACGTCACGTCACGTCACATGTCACATATTTCTTGGCTCTCGG	1140		
DB	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380		
QY	1141	GGATTTATTGTAGCATATCAGCTCTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA	1200		
DB	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400		
QY	1201	GAATTGATGAGACACACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCA	1260		
DB	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420		
QY	1261	ATTTTATAAGAAAATGCTACCTCATTTGCACGATTGCGCGAGATATTTCGAAATGCACCA	1320		
DB	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440		
QY	1321	AATGAAAATAACATTAGACACACACAGCATTTAAATCGACATCGTCTGTTCAATTC	1380		
DB	441	AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe	460		
QY	1381	GCCATGCTCCCAATTTTATGATCATATACAAACCCATATTAAATGATATGTTAGTAGG	1440		
DB	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480		
QY	1441	ATVCGCACAGCTTGTGGCAATTCAGAAATAGAGAACTGTTTATGGCACCAAGGATA	1500		
DB	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValIleThrPheGluGlyIle	500		
QY	1501	AGATTATCTTAGCGCTACACGAGTGCACATTAGGAAGGAGAGTGGCTGCAAGATG	1560		
DB	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet	520		
QY	1561	TTGGGGGATGCTGCTGCTATCGAGTGCATGCTATAGATCGGGAATCCGCTCACTTGG	1620		
DB	521	LeuGlyAspValAlaAlaValSerSerCysThrThrAlaIleAspAlaGluSerValThrLeu	540		
QY	1621	CAAAATCTTATCGAGTTTATCATCCACTTATACATGTTATAGCCGACCATTTGTTCTTA	1680		
DB	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu	560		
QY	1681	TTTTTCATATGGAGAAAACCAAGAAACATACAGGACCACTCGGTGAAAACACAGGTG	1740		
DB	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnGluLeu	580		
QY	1741	CTTCCAAACGCTAGAGGCTGAGAGCCATGCTCGGCTAAATCATCGTAGATATTTTCTGTTT	1800		
DB	581	LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe	600		
QY	1801	GGATCCGGTTATGCTTTATTTGAAACACTATAATTTTGTAAAGATGGTAGACGCTGCCGAT	1860		
DB	601	GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaLaasp	620		
QY	1861	ATACAGATTGTAGACATTTGTGAGCTTAATCTAACCTGCTAGAACATCGGGAAT	1920		
DB	621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640		

Db 285 ProTrpArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAlaArg 304
Qy 637 TCCTGTTCCGATTTCATATTTTGCATAGCCCAATGCGCAGCATCGCGAATATCTCCA 696
Db 305 SerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrAlaSerPro 324
Qy 697 TTTTATGCTCTATCCCCACAGAGGCTGCCGCA-----GAACCATGGGATATCCCCAG 750
Db 325 PheTyrGlyLeu-----ArgAlaAlaAlaAlaGlyIleGluHisAsnSerTyrAlaGln 341
Qy 751 GATAAATTTCAAACTAGATAGCTATTTTTCATGATTTGGACAGGCTGCAAAAGCA 810
Db 342 AspSerPheArgGlnValGluGlyTyrArgProArgAspLeuAspSerLysLeuGlnAla 361
Qy 811 AGCTTCCAGTCAAGGCTACTTTCATCATACATACACTTCACAGTTGGTGGGACTGG 870
Db 362 GluGluProValThrLysAsnPheIleThrProHisValThrValSerTyrAsnTyr 381
Qy 871 GCTCCAAAACTACTCGTGTATGTTCAATGACTTAAGTGGAAAGAGCTGACTGAAATGTTG 930
Db 382 ThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeuVal 401
Qy 931 CGTGCAACAGTTAATGGGAGATACAGATTTATGGCCCGTGAACCTTCGGCAACGTTTATC 990
Db 402 ArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerSerThrPheIle 421
Qy 991 AGTAATACGACTGAGTTTATCCAAATCGCATCATATTTAGGACATGTATTAACCGGAG 1050
Db 422 SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu 441
Qy 1051 GCAGAGCAGCAATCGAGCAGATATTTAGGACAAAATATATACACAGTCAAGTTC 1110
Db 442 AlaLysGluAlaIleAspSerIleTyrLysGlnTyrGluSerThrHisValPheSer 461
Qy 1111 GGACATGTCATATTTCTGGCTCTCGGGGATTTATGTACATATCAGCCTGTCTCTA 1170
Db 462 GlyAspValGluTyrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMetLeu 481
Qy 1171 TCCAAATCCCTGCTCATATGACCTCAGACGATTTGATGAGACACAGGACCGGATGAG 1230
Db 482 SerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyrAsp 501
Qy 1231 ATGCTGACCTGGTAAACAATAGCATGCAATTTATAGAAAATGCTACCTCATGTGCA 1290
Db 502 LeuLysAsnLeuLeuAsnProAsn-----AlaAsnAsnAsnAsnThrThr 517
Qy 1291 CGATTGCGCGAGATATTCGAATGCACCAATAGAAAATTAACATTAGAC----- 1341
Db 518 ArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyValHis 537
Qy 1342 -----GACACCCACAGCT----- 1353
Db 538 ArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAsp 557
Qy 1354 -----ATTAATCGACATCGCTGTCTCAATTC 1380
Db 558 SerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThrSerSerIleGluPhe 577
Qy 1381 GCATGCTCCAAATTTCTTATGATCATATACAAACCCATATTAATGATATGTTTAGTAGG 1440
Db 578 AlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSerArg 597
Qy 1441 ATTGCCACAGCTTGTGGGAATTCAGAAATAGAGAACTTGTTTTATGGCAGCAGGGATA 1500
Db 598 IleAlaThrAlaTyrCysThrLeuGlnAsnLysGluArgThrLeuTrpAsnGluMetVal 617
Qy 1501 AGATTATCTAGGCTACAGCGAGTCCACATTTAGGAAGAGAGTGGCTGCAAGATG 1560
Db 618 LysIleAsnProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgVal 637
Qy 1561 TTGGGGGATGCTGCTGTATGTCAGGCTGACTGCTATAGATGCGGAATCCGTCACATTG 1620
Db 638 LeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly---AsnValTyrLeu 656

Qy 1621 CAAAATTTATGCGAGTTATCACATCCACTAATACATGTTATATAGCCGACCATTTGTTCTA 1680
Db 657 GlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProValThr 675
Qy 1681 TTTTTCATAT-----GGAGAAAACCAAGGAACATACAGGACAACTCGGTGAAAC 1731
Db 676 PheThrIleThrLysAsnAlaAsnArgGlySerIleGluGlyGlnLeuGluGlu 695
Qy 1732 AACGAGTTGCTTCCACAGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCATGATAT 1791
Db 696 AsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArgTyr 715
Qy 1792 TTTTCTGTTTGGATCCCGTTATGCTTTTAAAGAACTATAATTTCTTAAAGTGGTAGAC 1851
Db 716 PheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThrPheValArgLysValPro 735
Qy 1852 GCTGCCGATATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAACCCCTGCTAGAGAT 1911
Db 736 ProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAsp 755
Qy 1912 CGGGAATTTTGCCTTTATCCGTTTACACAAAAGAGAGTTCGGTGATGTTGGTATTG 1971
Db 756 ArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeuLeu 775
Qy 1972 GATTATGCAAGATAGTCTCGCGCAATCACTACATGAACCTTAAATTTTATGACATAAAC 2031
Db 776 AspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeuArgPheTyrAspIleAsp 795
Qy 2032 AAAGTAATAGAGTGGATACAAAT-----AACCCAAATTAATGCGCAAAACC--- 2127
Db 796 SerValValAsnValAspAsnThrAlaValIleMetGlnGlyIleAlaSerPhePheLys 815
Qy 2062 GGGCTG-----CAGGAATTCGGCTGTATGCTCTATTACTTAAATAAAGATGCA--- 2106
Db 816 GlyLeuGlyLysValGlyGluAlaValGlyThrLeuValLeuAlaAlaGlyAlaVal 835
Qy 2107 -----AACCCAAATTAATGCGCAAAACC--- 2127
Db 836 ValSerThrValSerGlyIleAlaSerPheLeuAsnAsnProPheGlyGlyLeuAlaIle 855
Qy 2128 -----CAA 2130
Db 856 GlyLeuLeuValIleAlaGlyLeuValAlaAlaPhePheAlaTyrArgTyrValMetGln 875
Qy 2131 TTAGAGCAGCGCGAATGGAGTTAAGATCTAATCAATGCTTAAAGCGATGACATTAGCT 2190
Db 876 IleArgSerAsnProMetLysAlaLeuTyrProIleThrThrLysAlaLeuLysAsnLys 895
Qy 2191 TCACATACAGACTATGCCAAGATTCGAAGCTAGTTTATCATCTGCTTATAGTGAAGTGA 2250
Db 896 AlaLysThrSerTyrGlyGlnAsnGluGluAspGlySerAspPheAspGluAlaLys 915
Qy 2251 ACAGTTAAC-----AATAACCTTAATGCAACATTAGAACAA 2286
Db 916 LeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAlaLeuGluLysGln 935
Qy 2287 CTAAATAATGGCT---AAAACATAATTTAGAAATCAGCCATCAACCAAGCTAATACGATAAA 2343
Db 936 GluLysLysAlaIleLysLysAsnSerGlyValGlyLeuIleAlaSerAsnValSerLys 955
Qy 2344 ACGACTTTTGATTAATGAACACCCCAAAATTTAGTTGAAGCATCAAGCACTAAAACCACT 2403
Db 956 LeuAlaLeuArgArgArgGlyProLysTyr-----ThrArg 967
Qy 2404 TTAGAACCAACCTGCTACT 2421
Db 968 LeuGlnGlnAsnAspThr 973
RESULT 3
VGLB_HSVEB
ID VGLB_HSVEB
AC P28922; STANDARD; PRT; 980 AA.

DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520, 10330;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=AB4P.
RC STRAIN=Kentucky D.
RX MEDLINE=92295566; PubMed=1318606;
RA Teiford E.A., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1";
RL Virology 189:304-316(1992).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Kentucky D.
RX MEDLINE=90219246; PubMed=2157895;
RA Guo P.X., Goebel S., Perkins M.E., Taylor J., Norton E., Allen G.,
RA Lanquet B., Desmettre P., Paoletti E.;
RT "Coexpression by vaccinia virus recombinants of equine herpesvirus 1
RT glycoproteins gp13 and gp14 results in potentiated immunity.";
RL J. Virol. 64:2399-2406(1990).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86664; AAB02468.1; -;
DR EMBL: M35145; AAA46087.1; -;
DR PIR: G36798; VGB6C.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 980
FT DOMAIN 87 852
FT TRANSMEM 853 870
FT DOMAIN 871 980
FT CARBOHYD 165 165
FT CARBOHYD 275 275
FT CARBOHYD 380 380
FT CARBOHYD 423 423
FT CARBOHYD 497 497
FT CARBOHYD 514 514
FT CARBOHYD 515 515
FT CARBOHYD 560 560
FT CARBOHYD 727 727
FT CARBOHYD 749 749
FT SEQUENCE 980 AA; 10980 MW; A6DDFA8CA5550FF5 CRC64;
SQ
Alignment Scores:
Pred. No.: 4,91e-91 Length: 980
Score: 1779.50 Matches: 383
Percent Similarity: 55.94% Conservative: 135
Best Local Similarity: 41.36% Mismatches: 267
Query Match: 31.41% Indels: 141
DB: 1 Gaps: 18
US-09-147-052-3 (1-3261) x VGLB_HSVBB (1-980)
QY 13 AGCGCGAAGTGCATATTTTCCTTAGTTATTTCTATATGTT----- 54
DB 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85

QY 55 -----ACGAACATCATCTCCGAGTAGTACCCCAAAATGTGACATCA----- 90
DB ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis 105
QY 91 -----AGAGAAGTTGTT 102
DB GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY 103 TCGAGCGTCCAGTTGTCTGAGGAAGAGTCTAGCTTTTATCTTTTGTCTCCCCACACGAGGT 162
DB ArgGluSerGlyIleLeuAlaGluAspGlyAspPheThrThrCysProProThrGly 145
QY 163 TCAACCGTGATCCGTCTAGAACCGCGCGAANAATGTCGCAACCTAGAAAACCCACCGAG 222
DB SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGGTGAGGAATCGCGATATATTAAAGAGATATCATGCCATATTAATTTAAAGTG 282
DB PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProThrLysPheArgAla 185
QY 283 AGCCTTTATTATAAAATATCATTCAGACGACGACATGACGGGAGCAGCATATAGACAG 342
DB AsnValTyrrLysAspIleValThrArgValTrpLysGlyTyrrSerHisThrSer 205
QY 343 ATCACTAATCGATATACAGTAGAGCCCGCTTCCATTGAGAGATACGGATCTTAATC 402
DB LeuSerAspArgTyrrAsnAspArgValProValSerValGluGluIlePheGlyLeuIle 225
QY 403 GACGCGAAGGAGATGCTCATCTAAGCAAGATACCTTAGAACCAATGTATGTGAA 462
DB AspSerLysGlyLysCysSerSerLysAlaGluTyrrLeuArgAspAsnMetHisHis 245
QY 463 GCGTTTGACAGGATGCGGGAGAAAACAAAGTACTTCTAAACCAATCAAAATCAACACG 522
DB AlaTyrrHisAspAspLysAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
QY 523 CCCGAATCTAGGCGATGGCACACGACTAATAGACG-----TATACCGTGTGGGATCA 576
DB ProGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyrrValGlyTrp---Met 284
QY 577 CCATGGATATATCGAACGGGAACCCCGTCAATTTGTATAGTAGAGAAATGATGCCCGC 636
DB ProTrpArgHisTyrrThrSerThrSerValAsnCysIleValGluValGluAlaArg 304
QY 637 TCTGTGTTTCCGTATTCATATTTGCAATGGCAATGGCGACATCCGCAACATATATCCA 696
DB SerValTyrrProTyrrAspSerPheAlaLeuSerThrGlyAspIleValTyrrAlaSerPro 324
QY 697 TTTTATGGTCTATCCCCACGAGGCTGCCGCA-----GAACCCATGGGATATCCCGAG 750
DB PheTyrrGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyrrAlaGln 341
QY 751 GATAATTTCAACAACACTAGTACGTATTTTCAATGGATTTGACAAAGCGTCGAAAGCA 810
DB GluArgPheArgGlnValGluGlyTyrrArgProArgAspLeuAspSerLysLeuGlnAla 361
QY 811 AGCCTTCCAGTCAAGCGTAACCTTCTCATCATCATCATCATCATCATCATCATCATCAT 870
DB GluGluProValThrLysAsnPheIleThrThrProHisValThrValSerTrpAsnTrp 381
QY 871 GCTCCAAAACACTACTCGTGTATGTTCAATGACTAGTGGAAAGAGGTACTGAAATGTTG 930
DB ThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeuVal 401
QY 931 CGTGCACACAGTTAATGGGAGATACAGATTTATGGCCCGTGAACCTTCGCGACAGTTATC 990
DB ArgAspGluPheArgGlySerTyrrArgPheThrIleArgSerIleSerSerThrPheIle 421
QY 991 AGTAATACGACGTGAGTTTGTATCCAAATCGCATCATATATTAGGCAATGTATTAAACCGAG 1050
DB SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu 441

```

QY 1051 GCAGAAGCAGCAATCCAGCAGATATTAGCAAAATATAATGACAGTCACGCTCAAGGTT 1110
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 442 AlaLysGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPheSer 461
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1111 GCACATGTACATATTTCTTGCTCTCGGGGATTATTGTAGCATATCAGCCTGTTCTA 1170
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 462 GlnAspValGluTyrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgPrometLeu 481
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1171 TCCAAATCCCTGGCTCATATGTACCTCAGAGNAATGTAGCAGAGAACAGGACCATGAG 1230
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 482 SerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyrAsp 501
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1231 ATGCTCGACCTGGTAAACAATAAGCATGCAATTTATAAGAAAAATGCTACCTCATGTC 1290
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 502 LeuLysAsnLeuLeuAsnProAsn -----AlaAsnAsnAsnAsnThrThr 517
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1291 CGATTGCGCGGAGATATTGCAATGACCAATAGAAAAATAACATTAGAC----- 1341
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 518 ArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyValHis 537
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1342 -----GACACCACAGCT----- 1353
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 538 ArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAsp 557
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1354 -----ATTAATACGACATCGTCTGTCAATTC 1380
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 558 SerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThrThrSerSerIleGluPhe 577
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1381 GCCATGCTCCCAATTTCTTATGATCATATACAAACCCATATATATATATGTTAGTAGG 1440
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 578 AlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSerArg 597
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1441 ATTGCCACAGCTTGGTCGAATTCAGAAATAGAGAACTTGTATTGCGCAGAGGATA 1500
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 598 IleAlaThrAlaTyrCysThrLeuGlnAsnLysGluArgThrLeuTyrAsnGluMetVal 617
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1501 AAGATTAACTACCGCTACAGCAGTGCACATTAGAAGAGAGTGGCTGCAAGATG 1560
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 618 LysIleAsnProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgVal 637
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1561 TTGGGGATGTCGCTGCTATCGCTGATCGCTATAGTCCGGAATCCGTCACATTG 1620
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 638 LeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly-----AsnValTyrLeu 656
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1621 CAAATTTCTATGCGAGTTATCACATCCACTAATACATGTTATAGCCGACCATGTTCTA 1680
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 657 GlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProProValThr 675
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1681 TTTTTCATAT-----GGAGAAACCAAGGAACATACAGGACAACTCGGTGAAAC 1731
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 676 PheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGlyGlnLeuGlyGluGlu 695
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1732 AACGAGTTGCTTCCACGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCGTAGAT 1791
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 696 AsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArgTyr 715
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1792 TTTCTGTTGGATCCGGTTATGCTTTATTGAAAACATAATTTGTTAGATGGTAGAG 1851
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 716 PheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThrPheValArgLysValPro 735
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1852 GCTGCCGAVATACAGATTGTCACATTTGTCGACCTTAATCTAACCTGCTAGAGAT 1911
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 736 ProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAsp 755
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1912 CGGGAATTTGCTTATCCGTTTACAAAAAGAGAGTTGGCTGATGTTGGTGTATG 1971
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 756 ArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeuLeu 775
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 1972 GATTATGCAAGTAGCTCGCCGCAATCAACTACATGAATTAATTTATGACATAAAC 2031
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 776 AspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeuAlaArgPheTyrAspIleAsp 795
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2032 AAGTAATAGAGTGGATACAAAT-----TACGCG 2061
    ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|

```

```

Db 796 SerValValAsnValAspAsnThrAlaValIleMetGlnGlyIleAlaSerPhePheLys 815
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2062 GGGCTGCAGGAATTCGGC----- 2079
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 816 GlyLeuGlyLysValGlyGluAlaValGlyThrLeuValLeuGlyAlaAlaGlyAlaVal 835
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2080 -----TGTATGCTCTATTACTAAAAAAGATGCAAAACCCAAATATATGCGCAAAACC--- 2127
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 836 ValSerThrValSerGlyIleAlaSerPheLeuAsnAsnPropheGlyGlyLeuAlaIle 855
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2128 -----CAA 2130
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 856 GlyLeuLeuValIleAlaGlyLeuValAlaAlaPhePheAlaTyrArgTyrValMetGln 875
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2131 TTAGAAGCAGCGCAATGGAGTTAACAGATCTAATCAATGCTAAGCGATGACATTAGCT 2190
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 876 IleArgSerAsnPrometLysAlaLeuTyrProIleThrThrLysAlaLeuLysAsnLys 895
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2191 TCACACAAGACTATGCCAAGATTGAAGTAGTTATCATCTGCTTATCATCTGCTTATGAAGCTGAA 2250
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 896 AlaLysThrSerTyrGlyGlnAsnGluGluAspGlySerAspPheAspGluAlaLys 915
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2251 ACAGTTAAC-----AATAACCTTATGCAACATTAGAACAA 2286
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 916 LeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAlaLeuGluLysGln 935
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2287 CTAAAAATGGCT---AAAACATAATTAGAAATCAGCCATCAACCAAGCTAATACGGATAAA 2343
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 936 GluLysLysAlaIleLysLysAsnSerGlyValGlyLeuIleAlaSerAsnValSerLys 955
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2344 ACGACTTTTCATATGAACACCCCAATTTAGTTGAAGCATACAAAGCATTAAACACCAT 2403
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 956 LeuAlaLeuArgArgArgGlyProLysTyr-----ThrArg 967
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 2404 TTAGAACAACGTCGTACT 2421
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 968 LeuGlnGlnAsnAspThr 973
    |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|

```

RESULT 4

```

VGLB_VZVD
ID VGLB_VZVD STANDARD; PRT; 868 AA.
AC P09257;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein II).
GN 31.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04370; CAA27914.1; -.
CC FIR; E27214; VGBE31.
CC InterPro; IPR000234; Glycoprot_B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD000693; Glycoprot_B; 1.

```

KW	Glycoprotein; Transmembrane; Signal.	
FT	SIGNAL	1 ?
FT	CHAIN	? 868 GLYCOPROTEIN B.
FT	DOMAIN	? 679 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	680 695 POTENTIAL.
FT	TRANSMEM	701 720 POTENTIAL.
FT	TRANSMEM	724 744 POTENTIAL.
FT	DOMAIN	745 868 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	868 AA; 58B0E134F64FF902 CRC64;
Alignment Scores:		
Pred. No.:	6.3e-91	Length: 868
Score:	1777.50	Matches: 375
Percent Similarity:	61.30%	Conservative: 143
Best Local Similarity:	44.38%	Mismatches: 259
Query Match:	31.38%	Indels: 69
DB:	1	Gaps: 16
US-09-147-052-3 (1-3261) x VGLB_VZVD (1-868)		
QY	91	AGAGAAAGTGTGTTTCGAGCGTCCAGTGTCTGAGGAAGAGCTACGTTTATCTTTGTCCTCC 150
DB	41	ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
QY	151	CCACCAAGTGGTCAACCGTGATCGCTAGAACCGCCGGAATAATGTCCCGAACCTAGA 210
DB	61	ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
QY	211	AAAGCCACGAGTGGGGTGAAGAAATCGCGATATTATTTAAAGAGAAATACAGTCCATAT 270
DB	81	LeuGlyLysAsnPhetrhrGluGlyIleAlaValValTyrLysGluAsnIleAlaIaTyr 100
QY	271	AAATTAAAGTACCGTTTATATAAAATATATCATTTACAGACGACATGGAGGGGACG 330
DB	101	LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaThrAlaGlySer 120
QY	331	ACATATAGACAGTCACTAATCGATATACAGATAGGACGCCCGTTTCCATTGGAAGATC 390
DB	121	SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
QY	391	ACGGATCTAATCGACGGCAAGAGAGTGTCTATCTTAAAGAGAGATACCTTAGAAAACAAT 450
DB	141	ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn 160
QY	451	GTATATGTTGAAGCGTTTGACAGGGATCGGGAGAAAACAAGTACTTCTTAAACCATCA 510
DB	161	HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
QY	511	AAATTCAACGCCCCCAATCTAGGGATGGCACACCACTAATGAGACGTATACCGTGTGG 570
DB	181	LysTyrAsnSerValGlySerLysAlaThrPheHisThrAsnAspThrTyrMetValAla 200
QY	571	GGATCACCATGGATATATCAACCGGAACCTCCGTCAATTGTATAGTAGAGGAATGGAT 630
DB	201	GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
QY	631	GCCCGTCTGTGTTCCGTATTCATATTTTGCATATGGCCCAATGGCCACATCGCGAACATA 690
DB	221	AlaArgSerIlePheProTyrAspSerPheGlyLeuSerThrGlyAspIleIleTyrMet 240
QY	691	TCTCCATTTTATGGTCTATCCCAACAGAGCGTCCGCGAAGACCCATGGATATCCCCAG 750
DB	241	SerProPhePheGlyLeuArg---AspGlyAlaTyrArgGluHisSerAsnTyrAlaMet 259
QY	751	GATAAATTTCAAAACACTAGATACGTATTTTCAATGGATTGTCAGAACGCTCGAAAAGCA 810
DB	260	AspArgPheHisGlnPheGluGlyTyrArgGlnArgAspLeuAspThrArgAlaLeuLeu 279

QY	811	AGCCTTCCAGTCAAGCGTAACCTTTCTCATCATCATCATCATCTTACAGTTGGTGGGACTGG 870
DB	280	Glu---ProAlaAlaArgAsnPheLeuValThrProHisLeuThrValGlyTyrAsnTyr 298
QY	871	GCTCCAAAACACTACTCGTGTATCTTCAATGACTAAGTGGGAAAGAGTGACTGAATGTTG 930
DB	299	LysProLysArgThrGluValCysSerLeuValTyrPargGluValGluAspValVal 318
QY	931	CGTGCACACAGTTAATGGGAGATACAGATTTATGCCCGCTGAACCTTTCCGCAACGTTTATC 990
DB	319	ArgAspGluTyrAlaHisAsnPheArgPheThrMetLysThrLeuSerThrPheIle 338
QY	991	AGTAATACGACTGAGTTTGATCCAAATCGCATCATATTAGGACAAATGATTAACACGGAG 1050
DB	339	SerGluThrAsnGluPheAsnLeuAsnGlnIleHisLeuSerGlnCysValLysGluGlu 358
QY	1051	GCAGAAAGCAGCAATCGAGCATATTAGGACAAAATATATACAGTCTACGTCACAGGTT 1110
DB	359	AlaArgAlaIleIleAsnArgIleTyrThrArgTyrAsnSerSerHisValArgThr 378
QY	1111	GGACATGTACAATATTTCTTGGCTCTCGGGGATTTATTGTAGCATATACAGCTGTCTTA 1170
DB	379	GlyAspIleGlnThrTyrLeuAlaArgGlyGlyPheValValPheGlnProLeuLeu 398
QY	1171	TCCAAATCCCTGGCTCATATGTACCTCAGAGAATTGATGAGAGACACAGACCGATGAG 1230
DB	399	SerAsnSerLeuAlaArgLeuTyrLeuGlnGluLeuValArgGlu----- 413
QY	1231	ATGCTCGACCTGGTAAACAATAAGCATGCATTTTAAAGAAAATGCTACTCTATTGTCA 1290
DB	414	-----AsnThrAsnHisSer---ProGlnLysHisProThrArgAsnThr 427
QY	1291	CGATTCGGCGAGATATTCGAAATGCACCAATAGAAAATAACATATTAGACGACACCACA 1350
DB	428	ArgSerArgArgSerVal-----ProValGluLeuArgAlaAsnArg 441
QY	1351	GCTATTAAATCGACATCGTCTGTTCATTCGCCATGCTCCCAATTTCTTTATGATCATATA 1410
DB	442	ThrIleThrThrThrSerValGluPheAlaMetLeuGlnPheThrTyrAspHisIle 461
QY	1411	CAAACCAVATTAATGATATGTTTAGTAGGATGCCACAGCTGGTGGCGAATGCCGAAT 1470
DB	462	GlnGluHisValAsnGluMetLeuAlaArgIleSerSerSerTyrCysGlnLeuGlnAsn 481
QY	1471	AGAGAACTTTCTTTTATGGCAGGAGGATAAAGATTAACTAGCTACAGCGAGTCCA 1530
DB	482	ArgGluArgAlaLeuTyrPsrGlyLeuPheProIleAsnProSerAlaLeuAlaSerThr 501
QY	1531	ACATTAGGAAGAGAGTGGCTGCAAGAGATGTTGGGGATGCTCGCTGTATCGAGCTGC 1590
DB	502	IleLeuAspGlnArgValLysAlaArgIleLeuGlyAspValIleSerValSerAsnCys 521
QY	1591	ACTGCTATAGATCGGNAATCC---GTCACATTTCCAAAATTTCTATGCGAGTTATCACATCC 1647
DB	522	ProGluLeuGlySerAspThrArgIleIleLeuGlnAsnSerMetArgValSerGlySer 541
QY	1648	ACTAATACATGTTATATACCGCACCATTTGTTCTATTTCATATGGAGAAAACCAAGGAAC 1707
DB	542	ThrThrArgCysTyrSerArgProLeuIleSerIleValSerLeuAsnGlySerGlyThr 561
QY	1708	ATACAGGGACAACCTCGGTGAAAACAACAGAGTGTCTTCCAAACGCTAGAGGTAGAGCCA 1767
DB	562	ValGluGlyGlnLeuGlyThrAspAsnGluLeuIleMetSerArgAspLeuLeuGluPro 581
QY	1768	TGCTCGCTAATCATCGTAGATATTCTCTGTTGGATCCCGTTATGCTTTATTATTGAAAC 1827
DB	582	CysValAlaAsnHisLysArgTyrPheLeuPheGlyHisHisTyrValTyrGluAsp 601
QY	1828	TATAAATTTCTTAAAGTGGTAGACGCTGCCGATATATACAGATTGCTACACATTTGTCGAG 1887
DB	602	TyrArgTyrValArgGluIleAlaValHisAspValGlyMetIleSerThrTyrValAsp 621


```
QY 286 CTTTATTATAAATATCATTCAGACGACGACATGCGCGGGACGACATATACACAGATC 345
Db 173 IletyrIyrLysasnValIleValThrValtrpSerGlySerThrYrAlaAlaIle 192
QY 346 ACTAATCGATATACAGATGAGCGCGCTTCCATTGAAGAGATCACGGATCAATCGAC 405
Db 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
QY 406 GGCAAAGGAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATGTATATGTTGAAGCG 465
Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232
QY 466 TTTTGACAGGATCGCGGAGAAAAACAAGTACTTCTAAACCACTCAAAATTCACACGCC 525
Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
QY 526 GAATCTAGGCGATGCGACGACTAATAGACCATATACCGTGTGGGATCACCATGGATA 585
Db 253 GlyThrArgGlyTrpHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
QY 586 TATCGAAGCGGAACCTCCGTCATTTGTATAGTAGAGAAATGATGCGCGCTCTGTGTT 645
Db 273 TyrHisThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292
QY 646 CCCTATTTCATTTTGCATGCGCAATGGCGACATCGCGAACATATCTCCATTTTATGGT 705
Db 293 ProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrMetSerProPheTyrGly 312
QY 706 CTATCCCCACGAGGTCGCGCAGAACCCATGGGATATCCCCAGGATATTTCAACAA 765
Db 313 LeuArg--GluGlyAlaHisGlyGluHisIleGlyTyrAlaProGlyArgPheGlnGln 331
QY 766 CTAGATAGCTATTTTCAATGAGTTGGACAGCGTCGAAACGAGCCCTCCAGTCAG 825
Db 332 ValGluHisTyrTyrProIleAspLeuAspSerArgLeuArgAlaSerGluSerValThr 351
QY 826 CGTAACCTTCTCATCATCATCACATTCACAGTTGGTGGGACCTGGGCTCCAAAACACTACT 885
Db 352 ArgAsnProLeuArgThrProHisPheThrValAlaTrpAspTrpAlaProLysThrArg 371
QY 886 CGTGTATGTTCAATGACTAGTGGAAAGAGGTGACTGAAATGTGGT---GCAACAGTT 942
Db 372 ArgValCysSerLeuAlaLysTrpArgGluAlaGluMetThrArgAspGluThrArg 391
QY 943 AATGGGAGATACAGATTTATGCGCGTGAACCTTTCGGCAACGTTTATCAGTATACGACT 1002
Db 392 AspGlySerPheArgPheThrSerArgAlaLeuGlyAlaSerPheValSerAspValThr 411
QY 1003 GAGTTTGATCCAAATGCATCATATTAGGCAATGTATTAAACCGCAGGACGAGACGCA 1062
Db 412 GlnLeuAspLeuGlnArgValHisLeuGlyAspCysValLeuArgGluAlaSerGluAla 431
QY 1063 ATCGACGAGATATTTAGGCAAAATATATATGACAGTCACGTCAGGTTGGA---CATGTA 1119
Db 432 IleAspAlaIleTyrArgArgTyrAsnSerThrHisValLeuAlaGlyAspArgPro 451
QY 1120 CAATATTTCTGCTCTCGGGGATTTATTTGTAGCATATACAGCTGTTCTATCCAAATCC 1179
Db 452 GluValTyrLeuAlaArgGlyPheValValAlaPheArgProLeuIleSerAsnGlu 471
QY 1180 CTGGCTCATATGTACATCAGAGATTTGATGAGACACACAGGACCGATGAGATGCTCGAC 1239
Db 472 LeuAlaGlnLeuTyrAlaArgGluLeu-----GluArgLeuGly 484
QY 1240 CTGGTAAACAATAAGCATGCAATTTATAGAAAAATGCTACCTCATTTGTCACGATTCGGG 1299
Db 485 LeuAlaGly-----ValValGlyProAlaAlaProAlaAlaAlaArgAlaArg 501
QY 1300 CGAGAT-----ATTGAAATGCACCAATAGA 1326
Db 502 ArgSerProGlyProAlaGlyThrProGluProProAlaValAsnGlyThrGlyHisLeu 521
QY 1327 AAATTAACATTAGACACACACAGCTATTAATATCAGATCGCTGTTCAATTCGCCATG 1386
Db 522 ArgIleThr-----ThrGlySerAlaGluPheAlaArg 532
QY 1387 CTCAATTTCTTTATGATCATACAAACCCATATTAAATGATATTTTAGTAGATGTC 1446
Db 533 LeuGlnPheThrTyrAspHisIleGlnAlaHisValAsnAspMetLeuGlyArgIleAla 552
QY 1447 ACAGCTTGGTGGCAATTCAGATAGAGAACTTGTGTTTATGGCACGAAGGATTAAGATT 1506
Db 553 AlaAlaTrpCysGluLeuGlnAsnLysAspArgThrLeuTrpSerGluMetSerArgLeu 572
QY 1507 AATCTAGCGCTACACGAGTGCACATTTAGGAGGAGAGCTGCTGCCAAAGATGTTGGGG 1566
Db 573 AsnProSerAlaValAlaThrAlaAlaLeuGlyGlnArgValSerAlaArgMetLeuGly 592
QY 1567 GATGTGCTGCTATCTATCGAGCTGCTGCTATAGATGCGGAATCCGTCACCTTGGCAAAAT 1626
Db 593 AspValMetAlaIleSerArgCysValGluValArgGly--GlyValTyrValGlnAsn 611
QY 1627 TCTATGGGATTTATCATCTCCACTATATACATGTTATAGCCGACCATTTGTTCTATTTC 1686
Db 612 SerMetArgValProGlyGluArgGlyThrCysTyrSerArgProLeuValThrPheGlu 631
QY 1687 TATGGAGAAAACCAAGCAACATACAGGAGCAACTCGGTCAAAACACAGATTCCTTCCA 1746
Db 632 His---AsnGlyThrGlyValIleGluGlyGlnLeuGlyAspAsnGlnLeuLeuIle 650
QY 1747 ACCTAGAGGCTGTAGAGCCATGCTCGCTTAATCATCTAGATATATTTTCTGTTGGATCC 1806
Db 651 SerArgAspLeuIleGluProCysThrGlyAsnHisArgArgTyrPheLysLeuGlySer 670
QY 1807 GGTATGCTTTATTTGAAACTATATATTTTGTAAAGATGTAGAGCTGCGCATATACAG 1866
Db 671 GlyTyrValTyrTyrGluAspTyrAsnTyrValArgMetValGluValPro-----Glu 688
QY 1867 ATTGCTAGCACATTTGCTGAGCTTAATCTAACCTGCTAGAGATCGGGAATTTTCCT 1926
Db 689 ThrIleSerThrArgValThrLeuAsnLeuThrLeuLeuGluAspArgGluPheLeuPro 708
QY 1927 TTAGCTGTTTACAAAGAAGAGTTCGCTGCTGTTGCTGTTATGATGATATTCAGAGAAT 1986
Db 709 LeuGluValTyrThrArgGluLeuLeuAlaAspThrGlyLeuLeuAspTyrSerGluIle 728
QY 1987 GCTCGCGCAATCACTACATGACTTAATTTTATGACATAACAAAGTAATAGAGTG 2046
Db 729 GlnArgArgAsnGlnLeuHisAlaLeuLysPheTyrAspIleAspArgValLysVal 748
QY 2047 GATACAAAT-----TACGCGGGGCTGCAGGAATTC 2076
Db 749 AspHisAsnValValLeuLeuArgGlyIleAlaAsnPhePheGlnGlyLeuGlyAspVal 768
QY 2077 GGCTGT-----ATGCTCTATTACTAAATAAAGATGCAACCCCAATAAT 2118
Db 769 GlyAlaAlaValGlyLysValValLeuGlyAlaThrGlyAlaValIleSerAlaValGly 788
QY 2119 GGCAAAACCAATTAGAAGCAGCGCAATGGATTAACAGATCTAATCAATGCTAAAGCG 2178
Db 789 GlyMetValSerPhe-----LeuSerAsnProPheGlyAlaLeuAla 802
QY 2179 ATGACATAGCTTCTACATACAGACTATGCCAAGATTCGAAGCTAGTCTATCATCTGCTTAT 2238
Db 803 IleGlyLeuLeuValLeu-----AlaGlyLeuValAlaAlaPhe---LeuAlaTyr 818
QY 2239 AGTGAAGCTGAAACAGTTAACTTAATGCAACATATAGAACAACTAAAAATGGCT 2298
Db 819 ArgHisIleSerArgLeuArgArgAsnProMetLysAlaLeuTyrProVal---ThrThr 837
QY 2299 AAAACTAATTTAGATCAGCCATCAACCAAGCTTAATACGGATAAACAGCTTTTGTATPAT 2358
Db 838 LysThrLeuLysGluAspGlyValAspGluValAspValAspGluAlaLysLeuAspGln 857
QY 2359 GAACACCAATTTAGTTGAGCATACAAAGCACTAAAAACCACTTTAGAACACAGCT 2415
```


Db 858 AlaArgAspMetile-----ArgTyrMetSerIleValSerAlaLeuGluGlnGln 874
 RESULT 6
 VGLB_HSVEL
 ID VGLB_HSVEL STANDARD; PRT; 979 AA.
 AC P25218;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GL14 OR 33.
 OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=10330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90236317; PubMed=1692002;
 RA Guo P.;
 RT "Characterization of the gene and an antigenic determinant of equine
 RT herpesvirus type-1 glycoprotein 14 with homology to gB-equivalent
 RT glycoproteins of other herpesviruses.";
 RL Gene 87:249-255(1990).
 CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34861; AAA46086.1; -;
 DR PIR; JH0109; JH0109.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR Prodom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 979
 FT DOMAIN 87 979 GLYCOPROTEIN B.
 FT TRANSMEM 852 869 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 870 979 POTENTIAL.
 FT CARBOHYD 165 165 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 979 AA; 110331 MW; 9A1986B791C5B36 CRC64;
 Alignment Scores:
 Pred. No.: 4,81e-90 Length: 979
 Score: 1761.50 Matches: 351
 Percent Similarity: 62.87% Conservative: 118
 Best Local Similarity: 47.05% Mismatches: 200
 Query Match: 31.09% Indels: 77
 DB: 1 Gaps: 12
 US-09-147-052-3 (1-3261) x VGLB_HSVEL (1-979)
 Qy 13 AGCGGAATGTCATATTTTCCCTATAGTATTCATATGTT----- 54
 Db 66 ArgValGlyThrIleValLeuAlaCysLeuLeuLeuPheGlySerCysValValArgAla 85
 Qy 55 -----ACGAACATCATCTCCGAGTACCACAAATGTGACATCA----- 90
 Db 86 ValProThrThrProSerProThrSerThrProThrSerMetSerThrHisSerHis 105

QY 91 -----AGAGAAGTGTGT 102
 Db 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
 QY 103 TCGAGCGTCCAGTTCTGTGAGGAGAGTCTAGCTTTTATTTTGTCTCCCGACAGTGGGT 162
 Db 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145
 QY 163 TCAACCGTCATCCGTCGTAGAACCGCGGGAATCTCCGAACTCCGAACTAGAAAACCCACCGAG 222
 Db 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
 QY 223 TGGGCTGAAGGAATCGCATATATTAAAGAGAATATCAGTCCCATATAAATTTAAAGTG 282
 Db 166 PheThrGluGlyIleAlaValIlePheGlyGluAsnIleAlaProTyrLysPheArgAla 185
 QY 283 AGCCTTTATATAAATATCATTCAGACGACATGGGCGGAGCAGCATATATAGACAG 342
 Db 186 AsnValTyrTyrLysAspIleValValThrArgValTyrLysGlyTyrSerHisThrSer 205
 QY 343 ATCATAATCATATACAGATAGGAGCGCCGTTTCCATTGAAGAGATCAGGATCTAATC 402
 Db 206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeuIle 225
 QY 403 GACGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAA 462
 Db 226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMetHisHis 245
 QY 463 GCGTTTGACAGGGATGCGGAGAAACAACTACTTCTTAAACCATCAAAATCAACACG 522
 Db 246 AlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
 QY 523 CCCGAATCTAGGCGATGCCACACACTAATCAGACG-----TATACCGTGTGGGATCA 576
 Db 266 ProGlyAlaArgAlaTyrGlnThrThrAsnAspThrThrSerTyrValGlyTyr---Met 284
 QY 577 CCATGGATATATCGAAGCGGAACCTCGCTCAATTTGTATAGTAGAGAAATGGATCCCGC 636
 Db 285 ProTyrArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAlaArg 304
 QY 637 TCTGTGTTCCGTATTCATATTTTGCATGCGCAATGCCAATGCCGACATCGGACATATCTCCA 696
 Db 305 SerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrAlaSerPro 324
 QY 697 TTTTATGCTATATCCCGACAGGCTGCGCA-----GAACCCATGGGATATCCCGAG 750
 Db 325 PheTyrGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyrAlaGln 341
 QY 751 GATAATTTCAACAACACTAGATAGTATTTTCAATGGATTGGCAACAGCTCGAAAGCA 810
 Db 342 GluArgPheArgGlnValGluGlyTyrArgProArgAspLeuAspSerLysLeuGlnAla 361
 QY 811 AGCCTTCCAGTCAAGCGTAACTTCTCATCATCATCATCTTCCACATCTGCGTGGGACTGG 870
 Db 362 GluGluProValThrLysAsnPheIleThrThrProHisValThrValSerTyrPheIle 381
 QY 871 GCTCCAAAACACTACTCGTGTATCTCAATGACTAAGTGGAAAGAGTGAAGTGAATGTTG 930
 Db 382 ThrGluLysLysValGluAlaCysThrLeuThrLysTyrLysGluValValAspGluLeuVal 401
 QY 931 CGTGCAACAGTTAATGGAGATACAGATTATGGCCCGCGTGAACCTTCGCGAAGCTTATC 990
 Db 402 ArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerTyrPheIle 421
 QY 991 AGTAATACGACTGAGTTTGTATCCAAATCGCATCATATTAGGACAATGTATTAACCGGAG 1050
 Db 422 SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu 441
 QY 1051 GCAGAGACGACATCCGAGATATATTAGACAAAATATATGACACAGTCAAGGT 1110
 Db 442 AlaLeuGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPheSer 461


```
QY 1111 GGACATGTACAAATATTTCTGGCTCTCGGGGATTTATTGTAGCATATCAGCCCTGTTCTA 1170
Db 462 GlyAspValGluTyrLeuAlaArgGlyLeuAlaPheArgProMetLeu 481
QY 1171 TCCAAATCCCTGGCTCATATGCTCAGAGAATTTAGTACAGAACACAGGACCGATGAG 1230
Db 482 SerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyrAsp 501
QY 1231 ATGCTCAGCTGGTAAACATAAGCATGCAATTTATAGAAAAATGCTACTCTATGTCAC 1290
Db 502 LeuLysAsnLeuLeuAsnProAsn-----AlaAsnAsnAsnAsnThrThr 517
QY 1291 CGATTGGCGGAGATATTCGAAATCGCAATAGAAAATAACATTAGAC----- 1341
Db 518 ArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyValHis 537
QY 1342 -----GACACACACAGCT----- 1353
Db 538 ArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAsp 557
QY 1354 -----ATTAAATCGACATCGTCTGTGTTCAATTC 1380
Db 558 SerSerAsnValThrAlaLysGlnLeuGluLeuLysThrThrSerSerIleGluPhe 577
QY 1381 GCCATGCTCCAAATTTCTTTATGATCATATACAAACCATATTAATGATATGTTAGTAGG 1440
Db 578 AlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSerArg 597
QY 1441 ATTCCACAGCTTGGTGGATTCAGAAATAGAGAACTTGTATTATGGCAGAGGAGATA 1500
Db 598 IleAlaThrAlaTyrCysProLeuGlnAsnLysGluArgProLeuTyrAsnGluMetVal 617
QY 1501 AAGATTATCTAGCGGTACAGCGAGTGCAACATAGGAGGAGGTGGTGCACAAAGATG 1560
Db 618 LysIleThrProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgVal 637
QY 1561 TTGGGGATGTCGCTGTATCCAGCTGCACTCTATAGATCGGAATCGTCACATTTG 1620
Db 638 LeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly----AsnValTyrLeu 656
QY 1621 CAAATTTCTATGCGAGTTATCACATCCACTAATACATGTTATGCCAGCACTGGTCTA 1680
Db 657 GlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProProValThr 675
QY 1681 TTTTCATAT-----GGAGAAACCAAGAAACATACAGGACAACTCGGTGAAAC 1731
Db 676 PheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGlyGlnLeuGlyGluGlu 695
QY 1732 AACGAGTTGCTCCACAGCTAGAGCTCTAGAGCCATGCTCGGCTAATCATCTAGATAT 1791
Db 696 AsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArgTyr 715
QY 1792 TTTCTGTTTGGATCCGGTTATGCTTTATTTGAAACATAATTTTGAAGATGGTAGAC 1851
Db 716 PheLysPheGlyLysGluTyrValTyrTyroGluAsnTyrThrPheValArgLysValPro 735
QY 1852 GCTGCCGATATACAGATGTGTAGACATTTCTCGAGCTTAATCTAACCCCTCTAGAAGAT 1911
Db 736 ProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAsp 755
QY 1912 CGGAAATTTTGCCTTTATCCGTTTACACAAAGAGAGTTCGGTGATGTTGGTGATTG 1971
Db 756 ArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeuLeu 775
QY 1972 GATTATGAGAGTAGTTCGCGCAATCAACTACATGAACATTAATTTTATGACATAAC 2031
Db 776 AspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeuArgPheTyrAspIleAsp 795
QY 2032 AAAGTAATAGAGTGGAT 2049
Db 796 SerValValAsnValAsp 801
```

RESULT 7

```
VGLB_HSVE4
ID VGLB_HSVE4 STANDARD; PRT; 919 AA.
AC P17472;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID-10333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89125704; PubMed-2915378;
RA Riggio M.P., Cullinane A.A., Onions D.E.;
RT "Identification and nucleotide sequence of the glycoprotein gb gene
of equine herpesvirus 4.";
RL J. Virol. 63:1123-1133(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M26171; AAA46106.1; ALT_INIT.
DR PIR; A31880; VGBEQH.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 919 GLYCOPROTEIN B.
FT DOMAIN 41 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 809 POTENTIAL.
FT DOMAIN 810 919 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 919 AA; 103710 MW; C77E48F26C37BC7B CRC64;

Alignment Scores:
Align. No.: 5,45e-90 Length: 919
Score: 1760.50 Matches: 369
Percent Similarity: 58.68% Conservative: 138
Best Local Similarity: 42.71% Mismatches: 256
Query Match: 31.08% Indels: 101
DB: 1 Gaps: 16

US-09-147-052-3 (1-3261) x VGLB_HSVE4 (1-919)
QY 58 AACTCATCTCCGAGTACC-----CAAAATGTG 84
Db 34 SerSerGlnProSerThrProAlaSerThrGlnSerAlaLysThrValAspGlnThrLeu 53
QY 85 ACATCAAGAGAAGTGTGTTTCGAGCGCTCCAGTGTCTGAGGAGAGCTCT----- 132
Db 54 LeuProThrGluThrProAspProLeuArgLeuAlaValArgGluSerGlyIleLeuAla 73
QY 133 -----ACGTTTATCTTTGTCCTCCACAGCGTGGTTCACCGCTGCTGCTAGAA 183
Db 133 -----ACGTTTATCTTTGTCCTCCACAGCGTGGTTCACCGCTGCTGCTAGAA 183
```

Db 74 GluAspGlyAspPheTyrThrCysProProThrGlySerThrValValArgIleGlu 93
 QY 184 CCGCGCGGAAATATCCCGAACCTAGAAAAGCCAGGAGTGGGTGAAGAAATCCGGATA 243
 Db 94 ProProArgSerCysProLysPheAspLeuGlyArgAsnPheThrGluGlyIleAlaVal 113
 QY 244 TTATTTAAAGAGATATACAGTCCATATAATTTAAAGTGAGCGCTTATTTATATAAAATATC 303
 Db 114 IlePheLysGluAsnIleAlaProTyrLysPheArgAlaAsnValTyrTyrLysAspIle 133
 QY 304 ATTCAGACGACGATAGCGGGGACGACATATAGACAGATCACTATCCATCATACATACAGAT 363
 Db 134 ValValThrLysValTyrLysGlyTyrSerHisThrSerLysSerAspArgTyrAsnAsp 153
 QY 364 AGGACGCCCTTCCATTAAGAGATACAGGATCTAATCGACGGCAAGAAAGATGCTCA 423
 Db 154 ArgValProValSerValGluGluIlePheThrLeuIleAspSerLysGlyLysCysSer 173
 QY 424 TCTAAAGCAAGATACCTTAGAAAACATGTATATCTTGAACGCTTGACAGGGATCGGGA 483
 Db 174 SerLysAlaGluTyrLeuArgAspAsnIleMetHisAlaTyrHisAspAspGluAsp 193
 QY 484 GAAACACAGTACTCTTAAACCATCAAAATCAACACGCCCGAATCTAGGGCATGGCAC 543
 Db 194 GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaArgAlaTrpGln 213
 QY 544 ACAGTAATGAGACG-----TATACCGTGTGGGATCACCATGGATATATCGAAGCGGA 597
 Db 214 ThrThrAsnAspThrThrSerTyrValGlyTrp---MetProTrpArgHisTyrThrSer 232
 QY 598 ACTCCGCTCAATGTATAGTAAAGAAATGGATGCCGCTCTGTGTTCGCTATTCATAT 657
 Db 233 ThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspSer 252
 QY 658 TTGCAATGCCAATGGGCATCGGACATATCTCCATTTATGGTCTATCCCCACCA 717
 Db 253 PheAlaLeuSerThrGlyAspIleValTyrThrSerProPheTyrGlyLeuArg---Ser 271
 QY 718 GAGGCTCGCGACAGCCATGGATATCCCGAGGATAATTCACAACTAGATAGCTAT 777
 Db 272 AlaAlaGlnLeuGluHisAsnSerTyrAlaGlnGluArgPheArgGlnValGluGlyTyr 291
 QY 778 TTTTCANTCGATTGGCAAGCGTCGAAAGCAAGCGCTTCCAGTCAAGCGTAACCTTCTC 837
 Db 292 GlnProArgAspLeuAspSerLysLeuGlnAlaGlyGluProValThrLysAsnPheIle 311
 QY 838 ATCATCATCATCTCACAGTGGGTGGGACTGGCTCCCAAACTACTCGTGTATGTCA 897
 Db 312 ThrThrProHisValThrValSerTrpAsnTrpThrGluLysLysIleGluAlaCysThr 331
 QY 898 ATGACTAGTGAAGAGGTGACTGAAATGTTCGTGCAACAGTAAATGGGAGATACAGA 957
 Db 332 LeuThrLysTrpLysGluValAspGluLeuValArgAspGluPheArgGlySerTyrArg 351
 QY 958 TTTATGGCCCGNGAATTCGCAACCTTTATCAGTATACAGTACGATTTGATCCAAAT 1017
 Db 352 PheThrIleArgSerIleSerThrPheIleSerAsnThrGlnPheLysLeuGlu 371
 QY 1018 CGCATCATATTAGGCAATATTAAACGCGAGGAGAGAGCAAGCAATCGACAGATATTT 1077
 Db 372 AspAlaProLeuThrAspCysValSerLysGluAlaLysAspAlaIleAspSerIleTyr 391
 QY 1078 AGGACAAAATATATGACAGTCAGCTCAAGGTTGGACATGTACAAATATTTCTTGGCTCTC 1137
 Db 392 ArgLysGlnTyrGluSerThrHisValPheSerGlyAspValGluPheTyrLeuAlaArg 411
 QY 1138 GGGGATTTATGTAGCATATACGCTTCTTCTATCCAAATCCCGGCTCATATGACTC 1197
 Db 412 GlyGlyPheLeuIleAlaPheArgProMetIleSerAsnGluLeuAlaArgLeuTyrIleu 431
 QY 1198 AGAANAATTGATGAGACACAGGACCGCATGAGATGCTGACCTGGTAAACAAATAGCAT 1257
 Db 432 AsnGluLeuValArgSerAsnArgThrTyrAspLeuLysAsnLeuLeu---AsnProAsn 450

QY 1258 GCAATTTATAAGAAAATGCTACTCTCA----- 1284
 Db 451 AlaAsnHisAsnThrAsnArgThrArgArgSerLeuLeuSerIleProGluProThrPro 470
 QY 1285 -----TTCTCAGATTGGCGGAGATATTCGAAAT 1314
 Db 471 ThrGlnGluSerLeuHisArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGlu 490
 QY 1315 GCACCAAAAT-----AGAAAATAAATATAGACGACACACAGCTATTAATATCG 1362
 Db 491 AlaAlaAsnSerThrAsnSerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThr 510
 QY 1363 ACATCTCTCTTCAATTCGCAATTCCTTATTCATCATATATACAAACCCATAT 1422
 Db 511 ThrSerSerIleGluPheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisVal 530
 QY 1423 AATGATATCTTTAGTAGGATGCCACAGCTTGGTGGCAATTCAGAAATAGAGAACTTGT 1482
 Db 531 AsnGluMetLeuSerArgIleAlaThrAlaTrpCysThrLeuGlnAsnLysGluArgThr 550
 QY 1483 TTATGGCAGCAAGGGATAAGATTAATCTAGCGGTACAGCGAGTGCAACATTAGGAAG 1542
 Db 551 LeuTrpAsnGluMetValLysValAsnProSerAlaIleValSerAlaThrLeuAspGlu 570
 QY 1543 AGAGTGGCTGCCAAGATGTTGGGGATGCTGCTGTATCTCAGCTGAGCTGCTATAGAT 1602
 Db 571 ArgValAlaAlaArgValLeuGlyAspValIleAlaIleThrHisCysValLysIleGlu 590
 QY 1603 CGCGAATCCGTCACCTTTGCAAAATCTATCGAGTTATCACATCCACTAAATCATGTAT 1662
 Db 591 Gly---AsnValTyrLeuGlnAsnSerMetArg---SerSerAspSerAsnThrCysTyr 608
 QY 1663 AGCCGACCATGCTTCTATTTTCATATGGAGAAAC-----CAGGAAACATACAG 1713
 Db 609 SerArgProProValThrPheThrIleThrLysAsnAlaAsnSerArgGlyThrIleGlu 628
 QY 1714 GGCAACTCGGTGAAACACAGAGTTGCTTCCACGCTAGAGCTGAGAGCATGCTCG 1773
 Db 629 GlyGlnLeuGlyGluGluAsnGluValTyrThrGluArgLysLeuIleGluProCysAla 648
 QY 1774 GCTAATCATCTGAGATATTTCTGTTTGGATCCGCTTATGCTTTATTTGAAACATTAAT 1833
 Db 649 IleAsnGlnLysArgTyrPheLysPheGlyLysGluTyrValTyrThrGluAsnTyrThr 668
 QY 1834 TTTGTTAAGATGGTAGACGCTGCCGATATACAGATTGTAGACATTTGTGCGAGCTTAAT 1893
 Db 669 TyrValArgLysValProProThrGluIleGluValIleSerThrTyrValGluLeuAsn 688
 QY 1894 CTAACTCTGCTAGAGATCGGAAATTTGCTTTATCCGTTTATCCGTTTACAAAAAGAGTTG 1953
 Db 689 LeuThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeu 708
 QY 1954 CGTGATGTTGTTGATTTGATTATCGAAGTATGATAGAGTGGAT----- 2049
 Db 709 GluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeu 728
 QY 2014 AATTTTATGACATAAACAAGTAATAGAGTGGAT----- 2049
 Db 729 ArgPheTyrAspIleLeuAspSerValValAsnValAspAsnThrAlaValIleMetGlnGly 748
 QY 2050 -----ACAAATTCGCGGGCTCCAGAAATTCGCTGTATGCTATTAATAAAAAAGAT 2103
 Db 749 IleAlaThrPhePheLysGlyLeuGlyLysValGly----- 760
 QY 2104 GCAACCCCAATATATGCCAAACCAATTAAGACGCGCGAATGGAGTTAACAGATCTA 2163
 Db 761 -----GluAlaValGlyThrLeuValLeuGlyAla 770
 QY 2164 ATCATGCTTAAGCGATGACATTAGTCTCACTACAGACTATGCTCAAG----- 2211
 Db 771 AlaGlyAlaValValSerThrValSerGlyIleAlaSerPheIleAsnAsnProPheGly 790

```

QY 2212 -----ATTGAAGCTAGTTATCATCTCTTATAGTGAAGCTGAA 2250
Db 791 GlyLeuAlaIleGlyLeuLeuValIleAlaGlyLeuValAlaAlaPheAlaTyArg 810
QY 2251 ACAGTT-----AACATAACCTTAATGCAACATTAGAACAACTAAAAATGGCT 2298
Db 811 TyrValMetGlnLeuArgSerAnProMetLysAlaLeuTyProIleThrThrArgSer 830
QY 2299 AAACATAATTAGAACCTCAACCAAGCTAATACGGATAAAACGACTTTTGATAAT 2358
Db 831 LeuLysAsnLysAlaLysAlaSerTyGlyGlnAsnAspAspAspThrSerAspPhe 850
QY 2359 GAACACCCAAATTTAGTGAAGCA-----TACAAAGCACTAAAAACCACT 2403
Db 851 AspGluAlaLysLeuGluGluAlaArgGluMetIleLysTyMetSerMetValSerAla 870
QY 2404 TTAGAACAACGT 2415
Db 871 LeuGluLysGln 874

RESULT 8
VGLB_HSVE1
ID VGLB_HSVE1 STANDARD; PRT; 980 AA.
AC P18050;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10327;
RN [1]
RX MEDLINE=89279217; PubMed=2543744;
RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
RA Woodworth L.M.;
RT Identification and nucleotide sequence of a gene in equine
RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
RT major envelope glycoprotein gB."
RL J. Gen. Virol. 70:383-394(1989).
CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00401; BAA00304.1; ALT_SEQ.
DR PIR; A31241; VGBE2H.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 980 GLYCOPROTEIN B.
FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 853 870 POTENTIAL..
FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```

SQ SEQUENCE 980 AA; 109931 MW; 94B69AF882FB6AC2 CRC64;
Alignment Scores:
Pred. No.: 6.09e-89 Length: 980
Score: 1741.50 Matches: 379
Percent Similarity: 55.56% Conservative: 136
Best Local Similarity: 40.88% Mismatches: 269
Query Match: 30.74% Indels: 143
DB: 1 Caps: 20
US-09-147-052-3 (1-3261) x VGLB_HSVE1 (1-980)
QY 13 AGCGGGAATTCGATATATTTTCCCTATAGTTATCTATATGCT----- 54
Db 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
QY 55 -----ACGAACCTCATCTCCGAGTAGTACCCAAATGTGACATCA----- 90
Db 86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105
QY 91 -----AGAGAAGTTGTT 102
Db 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY 103 TCGAGCGTCCAGTTGTCTGAGGAAGACTCTACGTTTATCTTCTCCCCACCAGTGGGT 162
Db 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyThrCysProProThrGly 145
QY 163 TCAACCGTGATCGCTCTAGAACCGCGGAAATGTCCCGAACCTAGAAAAGCCACCGAG 222
Db 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGTCGAAGGAATCGCGATATATTTAAAGAGAATATCATCATATAAATTTAAAGTG 282
Db 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyLysPheArgAla 185
QY 283 ACCTTTATATAAATATCATTCAGACGACGATGGAGCGGGGACCATATAGACAG 342
Db 186 AsnValTyTyLysAspIleValValThrArgValTyLysGlyTySerHisThrSer 205
QY 343 ATCACAATCATATACAGTAGGCGCGCTTCCATTGCAAGAGATCACGGATCTAATC 402
Db 206 LeuSerAspArgTyTyrAsnAspArgValProValSerValGluGluIlePheGlyLeu 225
QY 403 GACGGCAAGGAAGATGCTCATCTAAAGCAAGTACCTTAGAACATATGATGTTGAA 462
Db 226 AspSerLysGlyLysCysSerSerLysAlaGluTyLeuArgAspAsnIleMethHis 245
QY 463 GCCTTTTCACAGGATCGGGGAGAAACAAGTA---CTTCTAAACCATCAAAATTCAC 519
Db 246 AlaTyHisAspAspGluAspGluValGluLeuAspLysCysArgProSer---LeuGln 264
QY 520 ACGCCGCAATCTAGGCGATGGCACACGACTAATGAGACG-----TATACCGTGTGGGA 573
Db 265 LeuArgGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyValGlyTrp--- 283
QY 574 TCACCATGGATATATCGAACGGAACTCCGTCATATGTATAGTAGAGAAATGGATGCC 633
Db 284 MetProTyrArgHisTyThrSerThrSerValAsnCysIleValGluValGluAla 303
QY 634 CGCTCTGTGTTTCCGTATTCATATTTTGCATATGGCCNAATGGCGCATCGGACATATCT 693
Db 304 ArgSerValTyTyProTyAspSerPheAlaLeuSerThrGlyAspIleValTyAlaSer 323
QY 694 CCATTTTATGCTCTATCCCGACAGAGCGTCCGCA-----GAACCCATGGGATATCCC 747
Db 324 ProPheTyGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyAla 340
QY 748 CAGGATAATTTCAAACAACATAGTACGTATTTTCAATGGATTTGGCAAGCTCGCAAAA 807
Db 341 GlnGluArgPheArgGlnValGluGlyTyArgProArgAspLeuAspSerLysLeuGln 360
QY 808 GCAGCGCTTCCAGTCAAGCGCTAACTTTCTCATCATCATCATCATCATCATCATCATCAT 867

```

```

Db      361 AlaGluGluProValThrLysAsnPheIleThrProHisValThrValSerTrpAsn 380
QY      868 TGGGCTCCAAAACACTACTCGTGTATGTTCAATGACTAAGTGAAGAGGTGACTGAAAGT 927
Db      381 TrpThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeu 400
QY      928 TTGCGTGCACACAGTTAATCGGAGATACACAGATTATGCGCCGCGTGAACCTTCGCGCAAGCTTT 987
Db      401 ValArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerThrPhe 420
QY      988 ATCAGTAATACGACGAGTTGTTGATCCAAATCGCATCAVATTAGGCAATGATTAAACGC 1047
Db      421 IleSerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLys 440
QY      1048 GAGGAGAGACGACGAGATGAGAGATATTTAGGCAAAATATATGACGATCACCTCAAG 1107
Db      441 GluAlaLysGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPhe 460
QY      1108 GTTGACATGTACAATATTTCTGGCTCTCGGGGATTTATGTAGCATATCAGCTGTT 1167
Db      461 SerGlyAspValGluTyrLysLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMet 480
QY      1168 CTATCCAAATCCCTGCCTCATATGTACCTCAGAGAAATGATGAGACAGACAGGACCGAT 1227
Db      481 LeuSerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyr 500
QY      1228 GAGATGCTCGACCTGTAACAACTAAGCATGCAATTTATAGAAAAATGCTACCTCATG 1287
Db      501 AspleuLysAsnLeuAsnProAsn-----AalaAsnAsnAsnAsnThr 516
QY      1288 TCACGATTGCGCGGAGATATTCGAATGCACCAATAGCAAAATAGCAATAGAC----- 1341
Db      517 ThrArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyVal 536
QY      1342 -----GACACACAGCT----- 1353
Db      537 HisArgGluGlnIleLeuHisArgLeuHisLysArgAlaValAlaThrAlaGlyThr 556
QY      1354 -----ATTAATCGACATCGCTGTTCAA 1377
Db      557 AspSerSerAsnValThrAlaLysGlnLeuLeuIleLysThrThrSerSerIleGlu 576
QY      1378 TTGCGCATGCTCCAAATTTCTTTATGATCATATACAAACCCATTAATGATGATTTGAT 1437
Db      577 PheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSer 596
QY      1438 AGGATTGCGACAGCTTGTGCGAATTCAGATAGACAACTGTTTATGCGCAGCAAGGG 1497
Db      597 ArgIleAlaThrAlaIleTyrCysThrLeuGlnAsnLysGluArgThrLeuTrpAsnGluMet 616
QY      1498 ATAAGATTATCTAGCGCTACAGCGAGTGCAACATTAGGAGGAGAGTGGCTGCAAG 1557
Db      617 ValLysIleAsnProSerAlaIleValSerAlaThrLeuAspGluArgValAlaIleArg 636
QY      1558 ATGTTGGGGATGTCGCTGTATGAGCTGCTGCTATAGATCGGAATCCGTCACCT 1617
Db      637 ValLeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly---AsnValTyr 655
QY      1618 TTGCAAAATCTATGCGCAGTTATCAGATCCACATTAATACATGTTATATGCCACCATGCGTT 1677
Db      656 LeuGlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProVal 674
QY      1678 CTATTTTCATAT-----GGAGAAACCAAGGAACATACAGGAGCAACCTCGGTGA 1728
Db      675 ThrPheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGlyGlnLeuGlyGlu 694
QY      1729 AACACACGTTGTTCCACGCTAGAGCGTGTAGAGCCATGCTCGGCTAATCATCTAGTA 1788
Db      695 GluAsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArg 714
QY      1789 TATTTTCTGTTTGGATCCGGTTATGCTTATTTTGAACACTATATTTTGTAGATGGTA 1848

```

RESULT 9

```

VGLB_HSV2H STANDARD; PRT; 904 AA.
AC P08666; P89450;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87071654; PubMed=3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gb glycoprotein gene of HSV-2 and
comparison with the corresponding gene of HSV-1."

```


QY 1378 TTGCGCATGCTCCAAATTTCTTTATGATCATATATACAAACCCATATATATATATATATAGT 1437
 Db PheAlaArgLeuGlnPheThrTyrAsnHisIleGlnArgHisValAsnAspMetLeuGly 519
 QY 1438 AGGATTGCGCAGCTTGTGCGGAATTCAGAGATAGAACTGTTTATATGCGCAGCAGGG 1497
 Db ArgIleAlaValAlaTrpCysGluLeuGlnAsnHisGluLeuThrLeuTrpAsnGluAla 539
 QY 1498 ATAAGATTATCTAGCGGTACAGCGATGCGCAACATTTAGGAAGAGAGTGGCTGCAAG 1557
 Db ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgArgValSerAlaArg 559
 QY 1558 ATGTTGGGGATGTCGCTGCTGATCAGCTGCACTGCTATAGATGCGGAATCGTCACT 1617
 Db MetLeuGlyAspValMetAlaValSerThrCysValProValAlaProAspAsnValIle 579
 QY 1618 TTGCAAAATCTATGCGAGTTATACATCCACTTAATACATGTTATATACCGCACCATGTT 1677
 Db ValGlnAsnSerMetArgValSerSerArgProGlyThrCysTyrSerArgProLeuVal 599
 QY 1678 CTATTTCATATGAGAGAAACCAAGAAAC---ATACAGGACAACCTCGGTGAAACAAC 1734
 Db SerPheArgTyr---GluAspGlnGlyProLeuIleGluGlyGlnLeuGlyGluAsnAsn 618
 QY 1735 GAGTTGCTTCCACGCTAGAGGCTGTAGAGCCATGCTCGCTAATCATCGTATATATTT 1794
 Db GluLeuArgLeuThrArgAspAlaLeuGluProCysThrValGlyHisArgArgTyrPhe 638
 QY 1795 CTGTTGGATCCGGTTATGCTTTATTTGAAACTATAATTTTGTAAAGTGGTAGAGCT 1854
 Db IlePheGlyGlyGlyValTyrPheGluGluTyrAlaTyrSerHisGlnLeuSerArg 658
 QY 1855 GCGGATATACAGATTGCTAGCACATTTGCGAGCTTAATCTAACCTGCTAGAGATCGG 1914
 Db AlaAspValThrThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAspHis 678
 QY 1915 GAAATTTGCTTTATCCGTTTACACAAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCT 1974
 Db GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp 698
 QY 1975 TATGACAGATGCTGCGCGCAATCACTACATGAACTTAAATTTTATGACATAACAA 2034
 Db TyrThrGlnArgGlnArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr 718
 QY 2035 GTAATGAAGTGGATCAAAAT-----TAGCGGGCTCGCAGGAATTC----- 2076
 Db ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPheGluGly 738
 QY 2077 -----GGCTGTGCTCTATTACTATAAAGATGCA 2106
 Db MetGlyAspLeuGlyArgAlaValGlyLysValValMetGlyValValGlyValVal 758
 QY 2107 AACCCAAATTAAGGCAACCAATTAGAGCAGCGCATGAGTTACAGATCAATC 2166
 Db SerAlaValSerGlyValSerSerPhe-----MetSerAsnProPhe 772
 QY 2167 AATGCTAAAGCATGACATTAAGTTTCACTACAAAGACTATGCAAGATGAAGTAGTTA 2226
 Db GlyAlaLeuAlaValGlyLeuValLeu-----AlaGlyLeu 785
 QY 2227 TCATCTGCTTATAGTAGAGCTGAACAGTT-----ACAATAACCTTAAATGCA 2274
 Db ValAlaAlaPheAlaPheAlaPheArgTyrValLeuGlnLeuGlnArgAsnProMetLysAla 805
 QY 2275 -----ACATTAGAACAATAAATGCGTAAACTAATTAATTAAGATCAGCCATC 2322
 Db LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyValGlyGlyGlu 825
 QY 2323 AACCAAGCTAATACGATTAACACGCTTTTATATGATGAACACCAAAATTTAGTTGAACA 2382
 Db GlyGluGluGlyAlaGluGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
 QY 2383 -----TACAAGACCTAAACACCACTTTAGAA-----CAACGCTGCTACT 2421

Db 844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
 QY 2422 AACCTTGAAGTTTGTGCATCACTGCTTATAATCAAAATCCCAATATATTTAGTGGATCTA 2481
 Db ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys 883
 QY 2482 TACAATAAAGCT 2493
 Db 884 ArgAsnLysAla 887
 RESULT 10
 VGLB_HSV23
 ID VGLB_HSV23 STANDARD; PRT; 904 AA.
 AC P06763;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL27 OR GB2.
 OS Herpes simplex virus (type 2 / strain 333).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87112925; PubMed=3027364;
 RA Stuve L.L., Brown-Shimer S., Pachel C., Najarian R., Dina D.,
 RA Burke R.L.;
 RT "Structure and expression of the herpes simplex virus type 2
 glycoprotein gb gene."
 RL J. Virol. 61:326-335(1987).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: GH, GB, GC, GG, GL, AND GE.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL: M15118; AAA45837.1; -
 DR PIR: A26790; VGBEB2.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 22
 FT CHAIN 1 22 POTENTIAL.
 FT DOMAIN 23 904 GLYCOPROTEIN B.
 FT TRANSMEM 728 743 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 749 768 POTENTIAL.
 FT TRANSMEM 772 792 POTENTIAL.
 FT DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 904 AA; 100186 MW; A8B36F74FDBCB539 CRC64;
 Alignment Scores:
 Pred. No.: 1.38e-88 Length: 904
 Score: 1735.00 Matches: 362
 Percent Similarity: 59.83% Conservative: 143
 Best Local Similarity: 42.89% Mismatches: 269

Query Match:	30, 63%	Indels:	70
DB:	1	Gaps:	16
US-09-147-052-3 (1-3261) x VGLB_HSV23 (1-904)			
QY	58	AACTCATCTCCGAGTACCCAAAATGTCAGTCAAGAGAAAGTGTTCGAGCGTCCAGTGG	117
DB	82	AsnAlaThrValAlaAlaGlyHisAlaThrLeuArgAlaHisLeuArgGluLeuVal	101
QY	118	TCTGAGGAAGACTCTACGTTTATCTTTCTGTCCTCCACCAGTGGTTCACACCGTAPCCGT	177
DB	102	GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValValGln	121
QY	178	CTAGAACCCCGCGGAAAATGTCGCGAACCTAGAAAAGCCACCGAGTGGGTGAAGAAATC	237
DB	122	PheGluGlnProArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyIle	141
QY	238	GGGATATTATTAAAGAGAATATCACTCCATATAAATTTAAAGTACGCTTTATTATAAA	297
DB	142	AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys	161
QY	298	AATATCATTCACACGACGACATGGACGGGACGACATATAGACAGATCACTAATCGATAT	357
DB	162	AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe	181
QY	358	ACAGATAGGACCGCCGTTTCCATTGAAGAGATCACGGATCTAATTCAGCGGCAAGGAAGA	417
DB	182	GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaLysGlyVal	201
QY	418	TGCTCATCTAAAGCAGATACCTTAGAACAAATGTATATGTTCAACGCTTTCACAGGGAT	477
DB	202	CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgSp	221
QY	478	GGGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCGCCGAATCTAGGCA	537
DB	222	AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArgGly	241
QY	538	TGGCACACACTAATGAGAGCTATACCGTGTGGGGATCACCATGGATATATCGAACGGGA	597
DB	242	TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly	261
QY	598	ACCTCCGTCAATGTATAGAGAAATGGATCCCGCTCTGTGTTCCGTATTCATAT	657
DB	262	ThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspGlu	281
QY	658	TTTGCAATGGCAATGGCAGACTCGCGAACAATCTCCATTTATGGTCTATCCCAACCA	717
DB	282	PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg	300
QY	718	GAGGCTGCCGAGAACCCATGGATATCCCGAGGATAATTCAAACAACCTAGATCTAT	777
DB	301	GlySerHisThrGluHisThrSerTyrAlaAlaAspArgPheLysGlnValAspGlyPhe	320
QY	778	TTTTCAATGGATTGGACAGGCTCGAAAAGCAAGCTTCCTCAGTCAAGCGTAACTTCTC	837
DB	321	TyrAlaArgAspLeuThrThrLysAlaArgAlaThrSerProThrThrArgAsnLeuLeu	340
QY	838	ATCACATCACTTCACGTTGGTGGGACTGGGCTCCAAAACACTACTCTGTTATTTCA	897
DB	341	ThrThrProLysPheThrValAlaTrpAspTyrValProLysArgProAlaValCysThr	360
QY	898	ATGACTAAGTGGAAAGGTGACTGAAATGTTGCGTGCACACAGTTAATGGGAGATACAGA	957
DB	361	MetThrLysTrpGlnGluValAspGluMetLeuArgAlaGluTyrGlySerPheArg	380
QY	958	TTTATGGCCGCTGAACCTTCGCGCAACGTTTATCAGTAATACGACTCAGTTTGATCCAAAT	1017
DB	381	PheSerSerAspAlaIleSerThrThrPheThrThrAsnLeuThrGlnTyrSerLeuSer	400
QY	1018	CGCATCATATTAGGACAATGTATTAAACGCGAGCGAGAAAGCAGCAATCAGCAGATATT	1077
DB	401	ArgValAspLeuGlyAspCysIleGlyArgAspAlaArgGluAlaIleAspArgMetPhe	420
QY	1078	AGGACAAAATATATATACAGCTACGTCACAGTTGGACATGTACAATATTTCTTGGCTCTC	1137
DB	421	AlaArgLysTyrAsnAlaThrHisIleLysValGlyGlnProGlnTyrTyrLeuAlaThr	440
QY	1138	GGGGGATTATTGTAGCATATCAGCTGTCTTCAAAATCCCTGCTCATATGTCACCTC	1197
DB	441	GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnThrLeuAlaGluLeuTyrVal	460
QY	1198	AGAAATTTGATGAGACACACAGGACCGATGAGATGCTGCAGCTGTGAACAATAAGCAT	1257
DB	461	ArgGluTyrMetArgGluGln	470
QY	1258	GCATTTTAAAGAAAATGCTACCTCATTTGTCACAGTTGCGGCGACATATTCCAAATGCA	1317
DB	471	ProArgAsnAlaThrPro	482
QY	1318	CCAAATAGAAAATAACATTTAGACGACACACAGCTATTAAATCGACATCTCTGTTCAA	1377
DB	483	ProSerAlaAsnAlaSerValGlu	499
QY	1378	TTGCCATGCTCCAAATTTCTTTATGATCATATACAAACCATATTATGATATGTTAGT	1437
DB	500	PheAlaArgLeuGlnPheThrTyrAsnHisIleGlnArgHisValAsnAspMetLeuGly	519
QY	1438	AGGATTTGCCACACCTTTGGTGGCAATTCGAGAATAGAGAACTGTTTATTCGCAGGAGG	1497
DB	520	ArgIleAlaValAlaTrpCysGluLeuGlnAsnHisGluLeuThrLeuTyrPsnGluAla	539
QY	1498	ATAACATTAATCTAGCGCTACAGCGAGTGCACATTAGAAGAGAGAGTGCACAAAG	1557
DB	540	ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgArgValSerAlaArg	559
QY	1558	ATGTTGGGATGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1617
DB	560	MetLeuGlyAspValMetAlaValSerThrCysValProValAlaProAspAsnValIle	579
QY	1618	TTGCCAAATTCATGCGGATTATCATCCACTAATACATGTTTATAGCCGACCATTCGTT	1677
DB	580	ValGlnAsnSerMetArgValSerSerArgProGlyThrCysTyrSerArgProLeuVal	599
QY	1678	CTATTTTCATATGGAGAAAACCAAGGAAC--ATACAGGGACAACCTCGGTGAACAC	1734
DB	600	SerPheArgTyr--GluAspGlnGlyProLeuIleGluGlnLeuGlyGluAsnAsn	618
QY	1735	GAGTTGCTTCCAACTAGAGCTGTAGAGCCATGCTCGGCTAATCATCTAGATATTT	1794
DB	619	GluLeuArgLeuThrArgAspAlaLeuGluProCysThrValGlyHisArgArgTyrPhe	638
QY	1795	CTGTTTGGATCCGTTTATGTTTGAACATATAATTTTGTAAAGATGCTAGACGCT	1854
DB	639	IlePheGlyGlyTyrValTyrPheGluGluTyrAlaTyrSerHisGlnLeuSerArg	658
QY	1855	GCGATATACAGATTGCTAGCACATTTGTCAGCTTAATCTAACCTGCTAGAGATCGG	1914
DB	659	AlaAspValThrThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAspHis	678
QY	1915	GAAATTTTGGCTTTATCCGTTTACAAAAGAGAGTTGCGTGTGTTGTTGTTGTTGTT	1974
DB	679	GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp	698
QY	1975	TATGCAGAGTACCTCGCCCAATCACTACATCACTTAATTTTATGACATTAACAAA	2034
DB	699	TyrThrGluValGlnArgArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr	718
QY	2035	GTAATAGAGTGGATCAAAAT--TACGCGGGCTCGAGGAATTC-----	2076
DB	719	ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPheGluGly	738
QY	2077	GGGTGTATGTCTTACTTAAAGATGCA	2106
DB	739	MetGlyAspLeuGlyArgAlaValGlyLysValValMetGlyValValGlyValVal	758
QY	2107	AACCCAAATANTGCCAAACCCCAANTTAGAAGACGCGGAATGGAGTTTACAGATCTAATC	2166


```

Db 759 SerAlaValSerGlyValSerSerPhe-----MetSerAsnProphe 772
QY 2167 AATGCTAAAGCGATGACATTAAGCTTCACTACAGACTATGCCAAGATTGAAGCTAGTTTA 2226
Db 773 GlyAlaLeuAlaValGlyLeuLeuValLeu-----AlaGlyLeu 785
QY 2227 TCATCGCTATAGTAGTGAAGCTGAACAGTT-----AACAAATACCTTAATGCA 2274
Db 786 ValAlaAlaPhePheAlaPheArgTyrValLeuGlnLeuGlnArgAsnPrometLysAla 805
QY 2275 -----ACATAGAACAACTAAATAAGCTGCTAAACCTAAATTTAGAACCCATC 2322
Db 806 LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyValGlyGlyGlu 825
QY 2323 AACCAAGCTAATACCGTAATAAGAGCTTTGTATGATGAACACCCCAATTTAGTTGAAGCA 2382
Db 826 GlyGluGluGlyAlaGlyGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
QY 2383 -----TACAAAGCACTAAATAAACCTTTAGTA-----CAAGCTGCTACT 2421
Db 844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
QY 2422 AACCTTGAGGTTTGTCACTCACTGCTTATATCAATCAATTCGAATAATTTAGTGGATCTA 2481
Db 864 ArgLysLysGlyThrSerAlaLeuLeuSerLysValThrAsnMetValLeuArgLys 883
QY 2482 TACAATAAAGCT 2493
Db 884 ArgAsnLysAla 897

RESULT 11
VGLB_HSVBC
ID VGLB_HSVBC STANDARD; PRT; 932 AA.
AC P12640;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130) (Glycoprotein B).
GN GI OR UL27.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID:10323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88300884; PubMed=2841484;
RA Whitbeck J.C., Bello L.J., Lawrence W.C.;
RT "Comparison of the bovine herpesvirus 1 gI gene and the herpes
simplex virus type 1 gB gene."
RL J. Virol. 62:3319-3327(1988).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21474; AAA6055.1; -
DR EMBL; J78205; CAB01598.1; -
DR EMBL; AJ004801; CRA06106.1; -
DR PIR; A28877; VGBEBC.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 67
FT CHAIN 68 932 GLYCOPROTEIN I.

```

```

FT DOMAIN 68 758 759 827 932 105 153 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT TRANSMEM 759 827 932 105 153 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT DOMAIN 828 932 105 153 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 105 153 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 153 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 441 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 483 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 640 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
FT CARBOHYD 706 706 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
SQ SEQUENCE 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;

Alignment Scores:
Pred. No.: 2.16e-88 Length: 932
Score: 1731.50 Matches: 341
Percent Similarity: 65.11% Conservative: 105
Best Local Similarity: 49.78% Mismatches: 218
Query Match: 30.56% Indels: 21
Db: 1 Gaps: 7

```

```

US-09-147-052-3 (1-3261) x VGLB_HSVBC (1-932)
QY 52 GGTACGACTCTCCGAGTACCACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTC 111
Db 97 GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116
QY 112 CAGTTGCTGAGGAGAGTCTACCTTTTATCTTTGTCCTCCACCCAGTGGTTCACACCGTG 171
Db 117 GlnAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal 136
QY 172 ATCCGCTAGAACCGCGCGGAAAATGTCGCGAACTAGAACGACCGAGTGGGTGAA 231
Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
QY 232 GGAATCGCATATTTTAAAGAGAAATATCAGTCCATATAAATTAAGTACGCTTTAT 291
Db 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIle 175
QY 292 TATAAAATATCATTCAGACGACGACATGACGGGAGACATATACAGACATCAATTAAT 351
Db 176 TyrLysAsnValIleValThrThrThrTipAlaGlySerThrThrAlaAlaIleThrAsn 195
QY 352 CGATATACATAGGACCGCGCTTCCATTTGAGAGACATCAGGATCTTAATCGACGCAAA 411
Db 196 GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLysLys 215
QY 412 GGAAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATGTATATGTTGAAGCGTTGAC 471
Db 216 TipArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValValAlaPheAsp 235
QY 472 AGGATGCGGGAGAAAACAAAGTACTTCTTAAACCATCAAAATTCACACGCGCAATCT 531
Db 236 ArgAspAspProTyrGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 255
QY 532 AGGCGATGGCACACGACTAATGAGACGTATACCGTGTGGGATCCAGTATATATCGA 591
Db 256 ArgGlyTyrPheThrThrAspAspValTyrThrAlaLeuGlySerAlaGlyLysArg 275
QY 592 ACGGGAACCTCCGCTCAATTCATAGTAGAGAAATGATGCGCGCTCTGTTTCGCTAT 651
Db 276 ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 295
QY 652 TCATATTTTCAATGGCCCAATGGGACATCGCGAACATATCTCCATTTTATGCTCTATCC 711
Db 296 AspSerPheAlaLeuSerThrGlyAspIleIleTyrMetSerProPheTyrGlyLeuArg 315
QY 712 CCACCAAGAGCGTCCGCCACCAACCCATGGATATCCAGGATAATTTCAACAACTAGAT 771
Db 316 ---GluGlyAlaHisArgGluHisThrSerTyrSerProGluArgPheGlnGlnIleGlu 334
QY 772 AGCTATTTTCAATGGATTGGACAAAGCGTCCAAAGCAAGCTTCCAGTCAAGCGTAC 831
Db 335 GlyTyrTyrLysArgAspMetAlaThrGlyArgArgLeuLysGluProValSerArgAsn 354

```


QY 832 TTTCTCATCATCAGTCTTCTGGTGGGAGTGGCTCCAAACACTACTCGTGA 891
 Db 335 PheLeuAthrGlnHisValThrValAlaThrPaspTrpValProLysArgLysAsnVal 374
 QY 892 TGTTCATCACTAACTGAAAGAGTGTGCTGAAATGTTGCGTCAACAGTTAATGGGAGA 951
 Db 375 CysSerLeuAlaLysTrpArgGluAlaAspGluMetLeuArgAspGluSerArgGlyAsn 394
 QY 952 TACAGATTATGGCCCGGAACTTTCGGCAACGGTTTATCATGAPATACGACTGAGTTGAT 1011
 Db 395 PheArgPheThrAlaArgSerLeuSerAlaThrPheValSerAspSerHisThrPheAla 414
 QY 1012 CCAATTCGCATCATATTAGGACAATGTATTAAACCGGAGGAGAGCAACATCCAGCAG 1071
 Db 415 LeuGlnAsnValProLeuSerAspCysValIleGluAlaGluAlaValGluArg 434
 QY 1072 ATATTAGCAAAATATAATACAGTACAGTCAAGGTTGGACATGTACAAATATTTCTTG 1131
 Db 435 ValTyArgGluArgTyArgGlyThrHisValLeuSerGlySerLeuGluThrTyArgLeu 454
 QY 1132 GCTCTCGGGGATTTATGTAGCATATCATGCTGTTCATCCAAATCCCTGCTCATATG 1191
 Db 455 AlaArgGlyGlyPheValAlaPheArgPrometLeuSerAsnGluLeuAlaLysLeu 474
 QY 1192 TACCTCAGAGATTGATGAGACACAGGACCGATGAGATGCTCGACCTGGTAAACAT 1251
 Db 475 TyrLeuGlnGluLeuAlaArgSerAsnGlyThrLeuGluGlyLeuPheAlaAla 492
 QY 1252 AAGCATGCAATTTATAAGAAAATGCTACCTCATTTGTCACGATTCGGCGAGATATCGA 1311
 Db 493 -----AlaAlaProLysProGlyPro-----ArgArgAlaArgAlaAlaPro 507
 QY 1312 AATGACCAAAATAGAAAATAACATA-----GACGACACCCACA 1350
 Db 508 SerAlaProGlyGlyProGlyAlaAlaAsnGlyProAlaGlyAspGlyAlaGlyGly 527
 QY 1351 GCTATTAAATCAGCATGCTCTCAATTCGCATGCTCCCAATTTCTTATCATATATA 1410
 Db 528 ArgValThrValSerSerAlaGluPheAlaAlaLysLeuAsnProSerAlaAlaSerAla 587
 QY 1411 CAACCCATATTATATGATTTAGTAGGATGGCAGAGTTCGCAATTCGAGAT 1470
 Db 548 GlnAspHisValAsnThrMetPheSerArgLeuAlaThrSerTrpCysLeuLeuGlnAsn 567
 QY 1471 AGAGAACTTGTGTTTGGCAGCAGGATTAAGATTAATCTAGCTACACGAGTGA 1530
 Db 568 LysGluArgAlaLeuTrpAlaGluAlaLysLeuAsnProSerAlaAlaSerAla 587
 QY 1531 ACATTAGGAGAGAGTGGCTCAAGATGTTGGGGATGCTGCTGCTATCGAGCTGC 1590
 Db 588 AlaLeuAspArgArgAlaAlaAlaArgMetLeuGlyAspAlaMetAlaValThrTyrcys 607
 QY 1591 ACTGCTATAGATCGGAATCCGCTCACTTTTGCAAAATTTCTATCGAGTTATCATCCACT 1650
 Db 608 HisGluLeuGlyGluGlyArgValPheIleGluAsnSerMetArg-----AlaProGly 625
 QY 1651 AATACATGTTATAGCGGACCATGTTCTATTTTCATATGAGGAGAAACCAAGAACATA 1710
 Db 626 GlyValCysTySerArgProValSerPheAlaPheGlyAsnGluSerGluProVal 645
 QY 1711 CAGGACACATCGTGAACACACGAGTTCCTCCACCGCTAGAGCTCTAGAGCCATGC 1770
 Db 646 GluGlyGlnLeuGlyGluAsnGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 665
 QY 1771 TCGGCTAATCATCGTAGATATTCTGTTTGGATCGGTTATGCTTTATTTGAAACATAT 1830
 Db 666 ThrAlaAsnHisLysArgTyArgPheGlyAlaAspTyArgTyArgGluAsnTyArg 685
 QY 1831 AATTTGTTAGATGAGTACGCTGCCGATATACAGATTCGATGATTCGAGCTT 1890
 Db 686 AlaTyValArgArgValProLeuAlaGluLeuGluValIleSerThrPheValAspLeu 705
 QY 1891 AATCTAACCCCTCTAGAGATCGGGAATTTTGCCTTTATCGCTTTACACAAAGAGAG 1950

Db 706 AsnLeuThrValLeuGluAspArgGluPheLeuProLeuGluValTyThrArgAlaGlu 725
 QY 1951 TTTCGTGTGTTGGTGTATTGATTATGATGAGAGTGTGCTCGCGCAATCAACTACATGAA 2010
 Db 726 LeuAlaAspThrGlyLeuLeuAspTyArgSerGluIleGlnArgArgAsnGlnLeuHisGlu 745
 QY 2011 CTTAAATTTTATGACATAACAAAGATATAGAAAGTGGATACAAATACGCG----- 2061
 Db 746 LeuArgPheTyArgPheLeuAspArgValValLysThrAspGlyAsnMetAlaIleMetArg 765
 QY 2062 GGCCTGCAGGAATTC 2076
 Db 766 GlyLeuAlaAsnPhe 770

RESULT 12
 VGLB_HVSM
 ID VGLB_HVSM STANDARD; PRT; 920 AA.
 AC Q04464;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN UL27
 OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset
 OC herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae;
 OC NCBI_TaxID=10353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93228440; PubMed=8385913;
 RA Eberle R., Black D.;
 RT "Sequence analysis of herpes simplex virus gb gene homologs of two
 RT platyrrhine monkey alpha-herpesviruses.";
 RL Arch. Virol. 129:167-182(1993).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M95786; AAA43841.1;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 920 GLYCOPROTEIN B.
 FT DOMAIN 29 739 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 740 755 POTENTIAL.
 FT TRANSMEM 761 781 POTENTIAL.
 FT TRANSMEM 784 804 POTENTIAL.
 FT DOMAIN 805 920 POTENTIAL.
 FT CARBOHYD 98 98 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;

Alignment Scores: 2.78e-88 Length: 920
 Pred. No.:

Score:	1729.50	Matches:	332
Percent Similarity:	64.21%	Conservative:	95
Best Local Similarity:	49.92%	Mismatches:	209
Query Match:	30.53%	Indels:	29
DB:	1	Gaps:	4
US-09-147-052-3 (1-3261) x VGLB_HSVSM (1-920)			
QY	91	AGAGAAAGTTGTTTCGAGCGCTCCAGTGTCTGAGGAGAGTCTACGTTTATCTTTGTC	150
Db		: : : : :	
QY	109	ArgGluSerValArgGlnIleArgAlaGluAsnAlaThrSerMetPheTyrValCysPro	128
Db		: : : : :	
QY	151	CCACCAAGTGGTTCACCGTGTCTAGAACCGCGGAAATGTCGCCAACCTAGA	210
Db		: : : : :	
QY	129	ProProThrGlyAlaThrValGlnPheGluGluProArgProCysProAspValAla	148
Db		: : : : :	
QY	211	NAAGCCCGAGTGGGTGAGGAATCGGATATTATTAAGAGAATATCAGTCCATAT	270
Db		: : : : :	
QY	149	AlaGlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyr	168
Db		: : : : :	
QY	271	AAATTTAAAGTCAACCGTGTATTAATAAATATCATTCAGACGACGATGGACGGGACG	330
Db		: : : : :	
QY	169	LysPheThrAlaThrMetTyrTyrLysGluIleThrValThrGlnThrTrpGlnGlySer	188
Db		: : : : :	
QY	331	ACATATAGACAGATCACTAATCGATATACAGATAGGACGCCGTTTCCATTGAAGAGATC	390
Db		: : : : :	
QY	189	ArgTyrLeuGlnLeuThrGlyLeuTyrAsnAspArgAlaProValProPheGluGluIle	208
Db		: : : : :	
QY	391	ACGGATCAATTCGAGCGGCAAGGAGATGCTCATCTAAGCAAGATACCTTAGAACAAT	450
Db		: : : : :	
QY	209	ThrAspValIleAsnAlaLysGlyLeuCysArgSerAspValThrTyrValArgSerGln	228
Db		: : : : :	
QY	451	GTATATGTTGAAGCGTTTGACAGGATGCGGGAGAGAAACAAAGTACTTCTAAACCATCA	510
Db		: : : : :	
QY	229	ArgArgValThrAlaTyrAspArgAspLutrpGlyArgGluValLysLeuValProSer	248
Db		: : : : :	
QY	511	AAATTCACACGCGCGAATCTAGGCGATGGCACAGCTACTTGTAGTAGAGGAATGGAT	570
Db		: : : : :	
QY	249	LysThrSerThrProAsnSerArgGlyTrpTyrThrThrAspArgMetTyrAlaProAsn	268
Db		: : : : :	
QY	571	GGATCACCAGTATATCGAACCGGAACCTCCGTCATTTGTAGTAGAGGAATGGAT	630
Db		: : : : :	
QY	269	AlaHisAlaGlyPheTyrLysAlaGlyThrValAsnCysIleValGluGluValGlu	288
Db		: : : : :	
QY	631	CGCGCTCTGTGTTCCGTATTCATATTTTGCATGCGCCCAATGGCGACATCGCGAACA	690
Db		: : : : :	
QY	289	AlaArgSerAlaTyrProTyrSerAsnPheValLeuAlaThrGlyAspPheValTyrVal	308
Db		: : : : :	
QY	691	TCTCCATTATGCTCTATCCACAGAGCGTCCCGCAGAACCCATGGGATATCCCCAG	750
Db		: : : : :	
QY	309	SerProPheTyrGlyLeuGly---GluAspAlaHisArgGluTyrAsnAlaTyrSerAla	327
Db		: : : : :	
QY	751	GATAATTTCAACAACACTAGATAGCTATTTTCAATGGATTGGACAAGCGTCCGAAAGCA	810
Db		: : : : :	
QY	328	AspArgPheLysGlnValAspGlyPhePheProArgAspLeuAspSerGlyGluThrAla	347
Db		: : : : :	
QY	811	AGCCTTCAGTCAAGCGTAACTTCTCATCATACATCATCTACAGTTGGTGGGACTGG	870
Db		: : : : :	
QY	348	ProGluProValArgAsnLeuLeuThrProLysPheThrIleGlyTrpAspTrp	367
Db		: : : : :	
QY	871	GCTCCAAAACACTACTCGTGTATCTCAATGACTTAAGTGAAGAGGTGACTGAATGTTG	930
Db		: : : : :	
QY	368	LysProLysAspProSerValThrLysTrpGlnValGluGluMetMet	387
Db		: : : : :	
QY	931	CGTGCAACAGTTAATGGGAGATACAGATTTATGGCCCGTGAACCTTCGGCAACGTTATC	990
Db		: : : : :	
QY	388	ArgAlaGluTyrGlySerThrPheArgPheThrSerSerLeuSerAlaThrPheThr	407
Db		: : : : :	
QY	991	AGTAATACGACGTGTTTCATCCCAATCCCATCATATATTAGGACAATGATTAAACCGGAG	1050
Db		: : : : :	
QY	408	ThrAsnValThrGlnTyrProGlnArgIleGluLeuSerAspCysValAlaArgGlu	427
Db		: : : : :	
QY	1051	GCAGAGCAGCAATCGAGCAGATATTTAGGACAAAATATTAATGACAGTCACGTCAGGTT	1110
Db		: : : : :	

RESULT 13
VGLB_HSV2S

ID VGLB_HSV2S STANDARD; PRT; 885 AA.
 AC P24994;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL27.
 OS Herpes simplex virus (type 2 / strain SA8) (Simian agent 8).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B264;
 RX MEDLINE=91374035; PubMed=1895066;
 RA Borchers K., Weigelt W., Buhr H.-J., Ludwig H., Mankertz J.;
 RT "Conserved domains of glycoprotein B (gp) of the monkey virus, simian
 agent 8, identified by comparison with herpesvirus gBs.";
 RL J. Gen. Virol. 72:2299-2304(1991).
 CC -|- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -|- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: GH, GG, GC, GD, GI, AND GE.
 CC -|- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.
 CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56935; CAA40256.1; --
 DR InterPro: IPR000234; Glycoprot.B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR Prodom: PD00693; Glycoprot_B; 1.
 KW Glycoprotein; Glycomembrane; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 885
 FT DOMAIN 35 715 GLYCOPROTEIN B.
 FT TRANSMEM 716 731 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 737 756 POTENTIAL.
 FT TRANSMEM 760 780 POTENTIAL.
 FT DOMAIN 781 885 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 659 659 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 885 AA; 97811 MW; 39E495B329AB94E4 CRC64;

 Alignment Scores:
 Pred. No.: 1,86e-87 Length: 885
 Score: 1714.50 Matches: 329
 Percent Similarity: 64.08% Conservative: 108
 Best Local Similarity: 48.24% Mismatches: 210
 Query Match: 30.26% Indels: 35
 DB: 1 Gaps: 5

 US-09-147-052-3 (1-3261) x VGLB_HSV2S (1-885)
 QY 58 AACTCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGGCTCGAGTGTG 117
 |||||
 Db 68 AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluIleIysAla 87
 |||||
 QY 118 TCTGAGGAAGAGTCTACGTTTACTTTGTCCTCCGCCACGAGTGGTTCACCGTGATCCGT 177
 |||||
 Db 88 ArgAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValValGln 107
 |||||
 QY 178 CTAGACCGCGCGGAAATGTCCTACCTCATTTGTCACGATTCGCGGAGAGATATTCGAATGCA 1317

Db 108 PheGluGlnProArgProCysProArgAlaProAspGlyGlnAsnTyrThrGluGlyIle 127
 QY 238 CGCATATTATTAAAGAGAGATATCAGTCCATATAAATTTAAAGTACGACCTTTATTATATAA 297
 |||||
 Db 128 AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 147
 |||||
 QY 298 AATATCATTCAGACGACATGAGCGGGGAGGACATATATAGACAGATCAGTAAATCGATAT 357
 |||||
 Db 148 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 167
 |||||
 QY 358 ACAGATAGGACGCGGTTTCCATTCAAGAGATACCGGATCTAATCGCGGGAAGGAAGA 417
 |||||
 Db 168 GluAspArgAlaProValPropheGluValMetAspLysIleAsnAlaLysIleVal 187
 |||||
 QY 418 TGCATCTCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGCTTTGACAGGAT 477
 |||||
 Db 188 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHisArgAsp 207
 |||||
 QY 478 CGGGAGAAAACAAGTACTTCTAAACCAATCAAAATTCACACGCCCGCAATCTAGGCA 537
 |||||
 Db 208 AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly 227
 |||||
 QY 538 TGGCACACGACTAATGAGACGATATACCGTGTGGGATCACCATGGATATATCGAACGGA 597
 |||||
 Db 228 TrpHisThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly 247
 |||||
 QY 598 ACCTCGCGCAATTTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGTTATTCATAT 657
 |||||
 Db 248 ThrThrValAsnCysIleValGluValGluAlaArgSerValTyrProTyrAspGlu 267
 |||||
 QY 658 TTTGCAATGGCAATGGGACATCCGAAACATATCTCCATTTTATGTTATGCCCAACA 717
 |||||
 Db 268 PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg---Asp 286
 |||||
 QY 718 GAGGTCGCGGCAACCCATCGGATATCCCGAGGATATTCACACACTAGATAGCAT 777
 |||||
 Db 287 GlySerHisGlyGluHisThrAlaThrAlaAlaAspArgPheArgGlnValAspGlyTyr 306
 |||||
 QY 778 TTTTCAATGGATTTGGACAAGCGTGAAGCAAGCCCTTCAGCTCAACGCTAACCTTC 837
 |||||
 Db 307 TyrGluArgAspLeuSerThrGlyArgArgAlaAlaAlaProValThrArgAsnLeu 326
 |||||
 QY 838 ATCATCATCACTTCCAGCTTGGGAGGCTCCAAACTACTCGTGTATGTTCA 897
 |||||
 Db 327 ThrThrProLysPheThrValGlyTrpAspTrpAlaProLysArgProSerValCysThr 346
 |||||
 QY 898 ATGACTAAGTGAAGAGGTGACTGAAATGTTGCTGCAACAGTTAATGGGAGATACAGA 957
 |||||
 Db 347 LeuThrLysTrpArgGluValAspGluMetLeuArgAlaGluTyrGlyProSerPheArg 366
 |||||
 QY 958 TTTATGGCCGCTGAACCTTTCCGCAACGTTTATCATAGTACTAGTGTGATCCCAAT 1017
 |||||
 Db 367 PheSerSerAlaAlaLeuSerThrThrPheThrAlaAsnArgThrGluTyrAlaLeuSer 386
 |||||
 QY 1018 CGCATCATATTAGGACATATTAACCGGAGGAGCAAGCAAGCAATCGACGAGATATTT 1077
 |||||
 Db 387 ArgValAspLeuAlaAspCysValGlyArgGluAlaArgGluAlaValAspArgIlePhe 406
 |||||
 QY 1078 AGGACAAAATATATGACAGTCAAGCTCAAGGTTGGCATGTACAAATATTTCTTGTGCTCTC 1137
 |||||
 Db 407 LeuArgArgTyrAsnGlyThrHisValLysValGlyGlnValGlnTyrTyrLeuAlaThr 426
 |||||
 QY 1138 GGGGATTTATGTAGCATATCAGGCTGTCTTATCCAAATCCCTGGCTCATATGTACCTC 1197
 |||||
 Db 427 GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnAlaLeuValGluLeuTyrVal 446
 |||||
 QY 1198 AGAGATTTGATGAGAGACACAGGACCGATGAGATGCTGCGACCTGGTGAACAATAAGCAT 1257
 |||||
 Db 447 ArgGluLeuValArgGlu----- 452
 |||||
 QY 1258 GCAATTTATAAGAAAAATGCTACCTCATTTGTCACGATTCGCGGAGAGATATTCGAATGCA 1317
 |||||
 Db 453 -----GlnThrArgArgProAlaGlyGlyAsp 461
 |||||

QY		1318	CCAAATAGAAAAATAACA-----TTAGACGACACCACAGCT-----ATTAAA	1359
Dd		462	ProGlycylAlaAlaThrProGlyProSerValAspProSerValGluArgIleLys	481
QY		1360	TCCGACATCGTCTGTTCATTAATGCCGCATCTCCAATTTCTTTATGATCATATACAAACCCAT	1419
Dd		482	ThrThrSerSerValGluPheAlaArgLeuGlnPheThrTyrAspHisIleGlnArgHis	501
QY		1420	ATTAATTCATATGTTTAGTAGGATGGCACAGCTTGTCGCAANTTGCAGATACAGAACCTT	1479
Dd		502	ValAsnAspMetLeuGlyArgIleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeu	521
QY		1480	GTTTTATGGCAGCAAGGGATAAAGATTAATCTAGCGCTACAGCGAGTGCACATTAGGA	1539
Dd		522	ThrLeuTrpAsnGlnAlaArgArgLeuAsnProGlyAlaIleAlaSerAlaThrValGly	541
QY		1540	AGGAGAGTGGCTGCAAGATGTTGGGGATGTCGCTGCTATCGAGCTGCACGTGCTATA	1599
Dd		542	ArgArgValSerAlaArgMetLeuGlyAspValMetalValSerThrCysValProVal	561
QY		1600	GATCGGGAATCCGCTACTTTTGCAAAATTCATCGCGAGTTATCACATCCCATAATCATGCT	1659
Dd		562	AlaProAspAsnValIleMetGlnAsnSerIleGlyValAlaAlaArgProGlyThrCys	581
QY		1660	TATAGCCGACCATGGTCTCTATTTTCATATGGAGAAMACCAAGGAACATACAGGGACAA	1719
Dd		582	TyrSerArgProLeuValSerPheArgTyrGluAlaAspGlyProLeuValGluGlyGln	601
QY		1720	CTCGGTGAAACAACGAGTGTGCTCCACGCTAGAGGCTGTAGAGCCATGCTCGGCTAAT	1779
Dd		602	LeuGlyGluAspAsnGluIleArgLeuGluArgAspAlaLeuGluProCysThrValGly	621
QY		1780	CATCGTAGATATTTCTGTTTGGATCCGGTATGCTTTATTTGAAAACATATAATTTGTT	1839
Dd		622	HisArgArgTyrPheThrPheGlyAlaGlyTyrValTyrPheGluGluTyrAlaTyrSer	641
QY		1840	AAAGATGGTAGAGCGTCGCATATACAGATTGCTACACATTTGTTCGAGCTTAATCAACC	1899
Dd		642	HisGlnLeuGlyArgAlaAspValThrValSerThrPheIleAsnLeuAsnLeuThr	661
QY		1900	CTGCTAGAAGATCGGGNAATTTTGCTTTATTCCTGTTTACAAAAAGAGAGTTGCGTGAT	1959
Dd		662	MetLeuGluAspHisGluPheValProLeuGluValTyrThrArgGlnGluIleLysAsp	681
QY		1960	GTTTGCTGATTCGGATTATGCGAAGTACTCTCGCGCAATCACTACATGAACCTTAATTT	2019
Dd		682	SerGlyLeuLeuAspTyrThrGluValGlnArgAsnGlnLeuHisAlaLeuArgPhe	701
QY		2020	TATGACATAAACAAAGTAAATAGAAGTGCATACAAT-----TAGCGCGGCTGCAG	2070
Dd		702	AlaAspIleAspThrValIleLysAlaAspAlaHisAlaAlaLeuPheAlaGlyLeuTyr	721
QY		2071	GAATTC	2076
Dd		722	SerPhe	723

RESULT, T 14

US-09-147-052-3 (1-3261) x VGLB_HSVB2 (1-917)	QY	58	AACTCATCTCCGGAGTAGTACCCAAAATGTGACATCAACAGAGAGTTGTTTCAGGCGCTCGAGTTG	117
	Db	110	AsnAlaSerGluProAlaAspProAlaGluLeuArgAlaAspLeuArgGlyLeuLysGly	129
	QY	118	TCTGAGGAAGAGTCTACTGTTTATCTTTTCTCCCCACCAGTGGTTCACCGGTGATCGGT	177
	Db	130	SerSerAspAspProAsnPhetYrValCysProProThrGlyAlaThrValValArg	149
	QY	178	CTAGAACCGCCGCGAAATGTCCCGAACCTAGAAAGCCACCGAGTGGGTGAAGGAATC	237
	Db	150	LeuGluGluProArgProCysProGluLeuProLysGlyLeuAsnPhetThrGluGlyIle	169
	QY	238	CGGATATTATTAAAGAGAATATCAGTCCATATAATTTAAAGTGACGCTTTATTATAA	297


```
Db 514 ArgHisValAsnAspMetLeuGlyArgValAlaIleAlaTrpCysGluLeuGlnAsnHis 533
QY 1474 GAAC TTGTTTATGCGACAGGATAAGATTATCTAGCGCTACAGCGAGTGCAACA 1533
Db 534 GluLeuThrLeuTrpAsnGluAlaArgLysLeuAsnProAsnAlaIleAlaSerAlaThr 553
QY 1534 TTAGCAAGGAGTGGCTGCCAAGATGTGGGGGATGTCGCTGCTGTATCGAGCTGCCACT 1593
Db 554 ValGlyArgValSerAlaArgMetLeuGlyAspValMetAlaValSerThrCysVal 573
QY 1594 GCTATAGATGGGAATCGCTCACTTTGCAAAATCTATCGAGTTATCACATCCACTAAT 1653
Db 574 ProValAlaAlaAspAsnValIleValGlnAsnSerMetArgIleSerSerArgProGly 593
QY 1654 ACATGTTATAGCCGACCATTTGTTCTATTTTCATATGGAGAAACCAAGGAAC---ATA 1710
Db 594 AlaCysTyrSerArgProLeuValSerPheArgTyr---GluAspGlnGlyProLeuVal 612
QY 1711 CAGGACAACTCGGTGAAACACAGCTTGCTTCCACGCTAGAGCGCTGTAGAGCCATGC 1770
Db 613 GluGlyGlnLeuGlyGluAsnGlnLeuArgLeuThrArgAspAlaIleGluProCys 632
QY 1771 TCGGCTAATCATCGTGTAGATATTTCTGTTGATCCGGTATGCTTTATTTCAAACTAT 1830
Db 633 ThrValGlyHisArgArgTyrPheThrPheGlyGlyGlyTyrValTyrPheGluGluTyr 652
QY 1831 AATTTTGTAAAGATGGTAGCGCTGCCGATATACAGATTGTGACACATTTGTGAGGCTT 1890
Db 653 AlaTyrSerHisGlnLeuSerArgAlaAspIleThrThrValSerThrPheIleAspLeu 672
QY 1891 AATCTAACCTGCTAGAGATCGGGAATTTTGCCTTTATCGCTTTTACACAAAGAAGAG 1950
Db 673 AsnIleThrMetLeuGluAspHisGluPheValProLeuGluValTyrThrArgHisGlu 692
QY 1951 TTGCGTGTATTGGTGTATTGGATTATCGAAGTAGCTCGCCCAATCAACTACATGAA 2010
Db 693 IleLysAspSerGlyLeuLeuAspTyrThrGluValGlnArgArgAsnGlnLeuHisAsp 712
QY 2011 CTTAAATTTTATGACATAACAAAGTAAATAGAAGTGGATACAAAT-----TACGCG 2061
Db 713 LeuArgPheAlaAspIleAspThrValIleHisAlaAspAlaAsnAlaAlaMetPheAla 732
QY 2062 GGGCTGCAGGAATTC 2076
Db 733 GlyLeuGlyAlaPhe 737
```

Search completed: October 8, 2003, 17:56:44
Job time : 130.558 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:32 ; Search time 107.546 Seconds
(without alignments)
5832.060 Million cell updates/sec

Title: US-09-147-052-3

Perfect score: 5665

Sequence: 1 atgcacatttttagcgga.....gaccgggtacatttttataa 3261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dallop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09147052/runat_08102003_154341_29860/app_query.fasta_1.5980
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052 -CGN_1_1_287=runat_08102003_154341_29860 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3601	63.6	865	1	VGBERB glycoprotein B pre
2	1898.5	33.5	948	2	A56602 glycoprotein B hom
3	1779.5	31.4	980	1	VGBEC6 glycoprotein B pre
4	1777.5	31.4	868	1	VGBE31 glycoprotein B - h
5	1775	31.3	913	1	VGBEBS glycoprotein gii p
6	1761.5	31.1	979	2	JH0109 glycoprotein 14 pr
7	1760.5	31.1	919	1	VGBEQH glycoprotein B pre
8	1755.5	31.0	975	2	T42576 Probable envelope
9	1743	30.8	943	2	B48474 glycoprotein B - f
10	1741.5	30.7	980	1	VGBE2H glycoprotein B pre
11	1735	30.6	904	1	VGBEB2 glycoprotein B pre
12	1731.5	30.6	932	1	VGBEBC glycoprotein gii pr
13	1727.5	30.5	920	1	D48349 glycoprotein B pre
14	1714.5	30.3	885	1	VGBESA glycoprotein B pre

15	1714	30.3	904	1	VGBEK2 glycoprotein B pre
16	1710	30.2	917	1	VGBEBH glycoprotein B pre
17	1705.5	30.1	903	1	VGBEB1 glycoprotein B pre
18	1705.5	30.1	904	1	VGBEW7 glycoprotein B pre
19	1699.5	30.0	903	1	VGBEK1 glycoprotein B pre
20	1671.5	29.5	933	1	B48349 glycoprotein B pre
21	1604	28.3	928	1	VGBEBG glycoprotein gii pr
22	1337.5	23.6	883	1	VGBEIS glycoprotein B pre
23	1325	23.4	883	1	VGBEIL glycoprotein B pre
24	1322	23.3	873	2	S26690 glycoprotein B - i
25	831.5	14.7	702	2	S48754 major surface prot
26	778.5	13.7	907	1	VGBETE glycoprotein B pre
27	760.5	13.4	854	2	T03107 glycoprotein B - a
28	756	13.3	906	1	VGBEC1 glycoprotein B pre
29	754	13.3	702	2	S48753 major surface prot
30	750	13.2	830	1	A44047 glycoprotein B pre
31	745	13.2	830	2	T43999 glycoprotein B (im
32	744	13.1	650	2	S48751 major surface prot
33	743	13.1	649	2	S48752 major surface prot
34	741	13.1	822	2	T41941 glycoprotein B - h
35	740.5	13.1	874	2	S25530 glycoprotein B - b
36	740	13.1	830	1	B44047 glycoprotein B pre
37	740	13.1	830	2	T44186 probable glycoprot
38	738.5	13.0	831	1	VGBE6S glycoprotein B - h
39	737	13.0	928	1	VGBEBC glycoprotein B pre
40	713	12.6	647	2	A49218 hemagglutinin homo
41	709.5	12.5	857	1	QGBEIL glycoprotein B - h
42	700	12.4	807	2	T42924 glycoprotein B - a
43	695	12.3	808	1	VGBESM glycoprotein B pre
44	680	12.0	874	2	S55602 glycoprotein B - e
45	575.5	10.2	386	2	S48755 major surface prot

ALIGNMENTS

RESULT 1

VGBERB

glycoprotein B precursor - Marek's disease virus (strain RB1B)

C:Species: Marek's disease virus

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C:Accession: A32402; B32402

R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog

A:Reference number: A32402; MUID:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 <ROS>

A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BA02866.1; PID:g221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271;304-330 <ROS2>

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-865/Product: glycoprotein B #status predicted <GPB>

F:709-728/Domain: transmembrane #status predicted <TN1>

F:732-752/Domain: transmembrane #status predicted <TN2>

F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Alignment Scores:

Pred. No.: 7.7e-202 Length: 865

Score: 3601.00 Matches: 719

Percent Similarity: 87.81% Conservativeness: 16

Best Local Similarity: 85.90% Mismatches: 60

Query Match: 63.57% Indels: 42

DB: 1 Gaps: 8

US-09-147-052-3 (1-3261) x VGBERB (1-865)

OY 1 ATGCACATTATTAGCGGAATTGGCATATTTTCCTTATAGTATCTATGTTACGTACGAAC 60

DB 1 MethHistyrPheArgAsnCysilePheLeulleValleLeuLeuTyrGlyThrAsn 20

QY 61 TCATCTCGGAGTACCCAAATGTGCATCAAGAGAACTTGTTCGACGGTCCAGTTGCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCGCCACCACTGAGTGGTTCACCGTGATCCGCTTA 180
Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GACCCGCGGAAATGTCGCCAGCTCCGAACTAGAAAGCCACCGAGTGGGTGAAGAAATCGCG 240
Db 61 GluProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
QY 241 ATATTATTTAAAGAGATATACGTCCTATATAATTTAAAGTGAGCGCTTTATTATAAAAT 300
Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCAGACGACGATGACGGGACGACATATAGACAGATCACTAATCCATATACA 360
Db 101 IleIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGGACGCGCTTCCATTGAAGAGATCACCGATCTAATCGACGCGCAAGGAAGATGC 420
Db 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAACGCTTTCACAGGGATCGG 480
Db 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAspAla 160
QY 481 GGAGAAAACAAGTACTTCTTAAACCATCAAAATTCACACGCGCCGAATCTAGGGCATGG 540
Db 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACACGACTAATGAGACGTATACCGTGTGGGATCACCATGATATATGATCAACGGGAACC 600
Db 181 HisThrAsnGluThrTyrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCGCTCAATGTATAGTAGAAATGGATGCCGCTCTGTGTTCCGTATTCATATTTT 660
Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCATGCCCAATGGGACATCGGACATATCTCCATTTTATGGTCTATCCCAACAGAG 720
Db 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProGlu 240
QY 721 GCTGCCGAGAACCCATGGATATCCCAAGGATAATTCACAACTATTCACAACTAGCTATT 780
Db 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe 260
QY 781 TCAATGGATTTGGACAAGCGTCGAAAGCAAGCCCTCCAGTCAAGCGTAACTTTCTCATC 840
Db 261 SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle 280
QY 841 ACATCACACTTCACAGTTGGGTGGGACTGGGCTCCAAAACTACTCGTGTATGTTCAATG 900
Db 281 ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrArgValCysSerMet 300
QY 901 ACTAAGTGGAAAGAGTCACTCAAAATGTCGTGCAACAGTTTAATGGGAGATACAGATT 960
Db 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
QY 961 ATGGCCCGTGAATCTTCGGCAAGCTTTATCAGTAAATACGACTGAGTTTGATCCAAATCGC 1020
Db 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
QY 1021 ATCATATTAGGCAATGTATTAAGCGGACGACGAGCAGCAGCAATCGACCATATTAGG 1080
Db 341 IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
QY 1081 ACAAAATATAATGACAGTCACGTCACAGGTGGACATGTACATAATTTCTTGGCTCTCGGG 1140
Db 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380

QY 1141 GGATTTATTTAGCATATCAGCTGCTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA 1200
Db 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
QY 1201 GAATGTGATGAGACACAACAGGACCGATGAGATGCTCGACCTGTGTAACAAATAAGCATGCA 1260
Db 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATAAGAAAATCTACCTCATTTCTCACGATTCGGGGGAGATATTCGAAATGCACCA 1320
Db 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgAspIleArgAsnAlaPro 440
QY 1321 AATAGAAAATACATTTAGACGACACACAGCTATTAAATCGACATCGCTGTTCAATTC 1380
Db 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerValGlnPhe 460
QY 1381 GCCATGCTCCATTTCTTTATGATCATATACAAACCATATTAATGATATGTTTAGTAGG 1440
Db 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
QY 1441 ATTGCCACAGCTTGGTCCGAATTCAGAAATGCAGATAGAGAACTTGTTTATGGCACGAAGGATA 1500
Db 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle 500
QY 1501 AAGATTAACTCTAGCGCTACAGCGAGTGCACATAGGAGAGAGAGTGGCTGCAAGATG 1560
Db 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
QY 1561 TTGGGGGATGTCGCTGTATTCGAGCTGCTGCTATAGATGCGGAATCCGTCACATTG 1620
Db 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
QY 1621 CAAAATTTCTATCGAGTTATCATCCACTAATACATGTATATAGCGGACCATTTGCTTCTA 1680
Db 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTTCATATGGAGAAAACAAGAACATACAGGACAACTCGGTGCAAAACACAGATG 1740
Db 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTCCAAAGCTAGAGCTGTAGAGCATGCTCGGCTAATCATCGTAGATATTTCTGTTT 1800
Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe 600
QY 1801 GGATCCGTTATGCTTTTATTTGAAACTATATTTTGTAAAGATGCTAGACGCTCCGAT 1860
Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
QY 1861 ATACAGATTTGATGACATTTGTCAGCTTATCTAACCCCTGCTGCTAGAGATCGGAAAT 1920
Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640
QY 1921 TTGCTTTATCCGTTTACACAAAAGAGATGCTGCTGATGTTGGTGTATTCGATATGCA 1980
Db 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla 660
QY 1981 GAAGTAGCTCCCGCAATCAACTACATGAATTTAAATTTTATGACATAAACAAGTATA 2040
Db 661 GluValAlaArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
QY 2041 GAAGTGGATACAAATATACGCG- - - - -GGGCTGCGAGGAATTCGGCTGTATGCTATT 2091
Db 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu- - - - - 695
QY 2092 ACTAAAAAGATGCAAAACCCCAATTAATGGCCAAACCCCAATTA- - -GAAGCAGCGCAATG 2148
Db 696 - - - - -PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
QY 2149 GAGTTAAGATCTAATCAATGCTTAACCGGATGACATTAGCTTCACACAGATATGCC 2208
Db 709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
QY 2209 AAG- - - - -ATTGAGCTAGTTTATCATCTGCT 2235

```

Db      729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
Qy      2236 TATAGTGAAGCTGAACAGTTAAC-----AATAACCTTAATGCA----- 2274
Db      749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
Qy      2275 ---ACATTAGAACAACTAAATAATGGCTAAACTAATTAGTAACTCAGCCATCAACCAAGCT 2331
Db      769 MetThrThrGluValGluLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp 788
Qy      2332 AATACGGATAAACGACTTTGTGATAATGAACACCCAAATTTAGTTGAAGCA----- 2382
Db      789 AspleuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet 806
Qy      2383 -----TACAAAGCACTAAACACCACTTTAGAACACAGCTGCTACTACCTT 2427
Db      807 IleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu 823

RESULT 2
A:56602
glycoprotein B homolog precursor - feline herpesvirus 1
C:Species: feline herpesvirus 1
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999
C:Accession: A56602
R:Maeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Niihara, M.; Kai,
Arch. Virol. 127, 387-397, 1992
A:Title: Identification and nucleotide sequence of a gene in feline herpesvirus type 1 H
A:Reference number: A56602; MUID:93090104; PMID:1333759
A:Accession: A56602
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-948 <MAE>
A:Cross-references: GB:S49775; NID:g261094; PIDN:AA24381.1; PID:g261095
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:120002, NCBIP:120003)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Alignment Scores:
Pred. No.: 6.92e-103 Length: 948
Score: 1898.50 Matches: 390
Percent Similarity: 59.53% Conservative: 163
Best Local Similarity: 41.98% Mismatches: 263
Query Match: 33.51% Indels: 113
DB: 2 Gaps: 14

US-09-147-052-3 (1-3261) x A56602 (1-948)
Qy      7 TATTTTAGCGGAATTGCATATTT----- 30
Db      22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
Qy      31 -----TTCCTTATAGTT 42
Db      42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleThrPle 61
Qy      43 ATCTATAT----- 51
Db      62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
Qy      52 -----GGTAGCAACTCATCTCCGAGTAGCCAA 78
Db      82 ProArgArgThrValAlaThrProGluValGlyGlyThrProProLysProThrThrAsp 101
Qy      79 AATGTGACATCAAGAAAGTTGTTTCGAGCGTCCAGTTGCTGAGGAAGAG----- 129
Db      102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
Qy      130 ---TCTACGTTTATCTTTGTCCTCCACCAAGTGGGTTCAACCGGTGATCCGCTAGAACCG 186
Db      122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141

```

```

Qy      187 CCGGAAATAGTCCCGAACCTAGAAAAAGCCACCGAGTGGGTGAAGNAATCGCCATATTA 246
Db      142 ProArgAlaCysProAspTyrTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
Qy      247 TTTAAACAGAAATATCAGTCCCATATAAATTTAAAGTGAGCGCTTTATATATAAAATATCATTT 306
Db      162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
Qy      307 CAGACGACGACATCGGCGGACGACATATATAGACAGATCACTAATTCGATATACAGATAGG 366
Db      182 MetThrThrValTyrSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
Qy      367 ACGCCCTTTCATTTGAAGAGATCAGGATCATATCGACGCAAGAGATGCTCATCT 426
Db      202 ValProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSer 221
Qy      427 AAACAAAGATACCTTAGAAACAATGTATATGTTCAAGCGTTTGACAGGATGCGGGAA 486
Db      222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro 241
Qy      487 AAACAAGTACTTCTAAACCACTCAAAATTTCAACACGCCCGCAATCTAGGGCATGGCACACG 546
Db      242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTyrPheThr 261
Qy      547 ACTAATGAGAGCTATACCGTGGGATCACCATGGATATATATGAACGGGAACCTCCGCTC 606
Db      262 ThrAsnGluThrTyrTyrLysIleGlyAlaAlaGlyPheHisHisSerGlyThrSerVal 281
Qy      607 AATGTATAGTAGAGAAATGGATGCCGCTCTGTTTCGTTATCATATATTTGCAATG 666
Db      282 AsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIle 301
Qy      667 GCCAATGCGGACATCGCAACATATCTCCATTTTATGCTCTATCCCAACAGAGGCTGCC 726
Db      302 SerThrGlyAspValIleHisMetSerProPhePheGlyLeuArg---AspGlyAlaHis 320
Qy      727 GCACAAACCATGGGATATCCCAAGGATATTTTCAACACACTAGATAGCTATTTTCAATG 786
Db      321 ValGluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrTyrProIle 340
Qy      787 GATTTGACAAAGCGTCGAAAAAGCAAGCCTTCAGTCAAGCGTAACCTTCTCATCACATCA 846
Db      341 AspLeuAspThrArgLeuGlnLeuGlyAlaProValSerArgAsnPheLeuGluThrPro 360
Qy      847 CACTTCACAGTTGGTGGGACTGGGCTCCAAAAACTACTCGTGTATGTTCAATGACTAAG 906
Db      361 HisValThrValAlaTyrAsnTyrProLysCysGlyArgValCysThrLeuAlaLys 380
Qy      907 TGGAAAGAGGTGACTGAAATGTTGCGTGCAACAGTTAATGGAGATACAGATTTATGCCC 966
Db      381 TrpArgGluIleAspGluMetLeuArgAspGluTyrGlnGlySerTyrArgPheThrVal 400
Qy      967 CGTCAACTTCGCGCAAGCTTTATCAGTAATACGACTGAGTTTGATCCAAATCGCATCATA 1026
Db      401 LysThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArg 420
Qy      1027 TTAGGACAATGATTTAAACCGGAGAGCAAGCAACATCGACAGATATTTAGACAAAA 1086
Db      421 LeuGlyAspCysAlaThrLysGluAlaAlaGluAlaIleAspArgIleTyrLysSerLys 440
Qy      1087 TATATGACAGTCAAGTCAAGTTGGACATGTACAATATTCTTGGCTCTCGGGGATTT 1146
Db      441 TyrSerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPhe 460
Qy      1147 ATTGTAGCATATCAGCTGCTTCTATCCAAATCCCTGGCTCATATGTACCTCAGAGAAATG 1206
Db      461 LeuIleAlaPheArgProMetIleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeu 480
Qy      1207 ATGACAGACACAGACCGGATGAGATGCTCGACCTGGTGTAAACAATAAGCATGCAATTAT 1266
Db      481 AlaArgSerAsnArgThrValAspLeuSerAlaLeuLeuAsnProSerGlyGluThrVal 500
Qy      1267 AAAAAAATGCTACCTCATTTG-----TCAGATTTGGGGGAGATATTCGAAATGACACA 1320

```


Db 126 ArgGluSerGlyIleLeuAlaGluaspGlyAspPheTyrThrCysProProThrGly 145
QY 163 TCACCGTGCATCGGTGTAGAACCGCGCGGAAATGTCCGAACTAGAAAAGCCCGAG 222
Db 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGGTGAAGGATCCGATATATTAAAGAGAATATCAGTCCATATAAATTTAAAGTG 282
Db 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheArgAla 185
QY 283 AGCTTTATTATAAAATATCATTCAGACGACATGGCGGGGACGACATAGACAG 342
Db 186 AsnValTyrTyrLysAspIleValThrArgValTyrLysGlyTyrSerHisThrSer 205
QY 343 ATCACTAATCGATATACAGATAGGACCGCGGTTTCCATTGAAGAGATCAGGATCAATC 402
Db 206 LeuSerAspArgTyrAsnArgValProValSerValGluIlePheGlyLeuIle 225
QY 403 GACGGCAAGGAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATGTATGTGAA 462
Db 226 AspSerLysGlyLysCysSerLysAlaGluTyrLeuArgAspAsnIleMetHisHis 245
QY 463 GGGTTTACAGGATCGGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACG 522
Db 246 AlaTyrHisAspAspGluaspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
QY 523 CCGGAATCTAGGCATGGCACGAGCTAATGAGACG-----TATACCGTGGGGATCA 576
Db 266 ProGlyAlaAlaArgIleThrThrAsnAspThrThrSerTyrValGlyTrp---Met 284
QY 577 CCATGGATATACGACGGGACCTCCGTCATTTGATAGTAGAGAAATGGATCCCGC 636
Db 285 ProTrpArgHisTyrThrSerThrSerValAsnCysIleValGluValGluAlaArg 304
QY 637 TCTGTGTTTCCGTATTTCATATTTCGAATGGCAATGGCGACATCCGCAACATATCTCCA 696
Db 305 SerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrAlaSerPro 324
QY 697 TTTTATGGTCTATCCCGACAGAGGTGCCGA-----GACCCCATGGATATCCCGAG 750
Db 325 PheTyrGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyrAlaGln 341
QY 751 GATAATTTCAACAACATAGATAGTATTTCATGGATTTCGACAAAGCTCGAAAGCA 810
Db 342 GluArgPheArgGlnValGluGlyTyrArgProArgAspLeuAspSerLysLeuGlnAla 361
QY 811 AGCCTTCCAGTCAAGCGTAACCTTCTCATCATCACACTTCACAGTTGGGTGGGACTGG 870
Db 362 GluGluProValThrLysAsnPheIleThrThrProHisValThrValSerTrpAsnTrp 381
QY 871 GCTCCAAAACACTCGTGTATGTTCAATGACTAAGTGAAGAGGTGACTGAATGTTG 930
Db 382 ThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeuVal 401
QY 931 CGTGCAACAGTTAAATGGAGATACAGATTATGGCCCGTGAATTCCTGGCAAGTTTATC 990
Db 402 ArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerSerThrPheIle 421
QY 991 AGTATPACGACTGAGTTTGATCCAAATCGCATCATATTAGGCAATGTATTAACCGCGAG 1050
Db 422 SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu 441
QY 1051 GCAGAAGCAGCAATCAGCAGATATTTAGGCAAAATATATAGTACAGTCCAGGT 1110
Db 442 AlaLysGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPheSer 461
QY 1111 GGACATGTACAATATTTCTGCTCGCGGGATTTTGTAGCATATCAGCTGTCTTA 1170
Db 462 GlyAspValGluTyrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMetLeu 481
QY 1171 TCCAAATCCCTGGCTCATATGTACCTCAGAGATTCATGATGACAGACAAACAGGACGATGAG 1230
Db 482 SerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyrAsp 501

QY 1231 ATGCTCGACCTGGTAAACAATAAGCATGCAATTTTATAAGAAAAATGCTACCTCATTTGTCA 1290
Db 502 LeuLysAsnLeuLeuAsnProAsn-----AlaAsnAsnAsnAsnThrThr 517
QY 1291 CGATTCCGGCGGAGATATTGCAATGACCAATAGAAAATAGAAAATACATTTAGAC----- 1341
Db 518 ArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyValHis 537
QY 1342 -----GACACACACAGCT----- 1353
Db 538 ArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAsp 557
QY 1354 -----ATTAATTCGACATCGTGTGTCAATTC 1380
Db 558 SerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThrThrSerSerIleGluPhe 577
QY 1381 GCCATCTCCAATTTCTTATGATCATATACAAACCCATATTAATGATATGTTTAGTAGG 1440
Db 578 AlaMetGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSerArg 597
QY 1441 ATTGCCACAGCTTGGTCGAATTCAGAACTGTTTATAGGACAGAGGATA 1500
Db 598 IleAlaThrAlaTrpCysThrLeuGlnAsnLysGluArgThrLeuTrpAsnGluMetVal 617
QY 1501 AAGATTAACTTAGCCCTACAGCGAGTGCACAACTTAGGAGAGAGTGCCTGCAAAAGATG 1560
Db 618 LysIleAsnProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgVal 637
QY 1561 TTGGGGGATGCTGCTGCTATCGAGCTGCACCTGCTATAGATCGGGAATCCGTCACATTG 1620
Db 638 LeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly---AsnValTyrLeu 656
QY 1621 CAAAATTCATGCGAGTTATCATCATCCACTATATACATGTTATAGCCGACCATGGTCTTA 1680
Db 657 GlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProProValThr 675
QY 1681 TTTTTCATAT-----GGAGAAAACCAAGGAACATCAGGAGCAACTCGGTGAAAAC 1731
Db 676 PheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGlyGlnLeuGlyGluGlu 695
QY 1732 AACGACTGCTTCCACGCTAGAGGTGTAGAGCCATGCTCGCTAATCATCTAGATAT 1791
Db 696 AsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArgTyr 715
QY 1792 TTTCTGTTGGATCCGTTATGCTTATTGTAACATATAATTTTCTTAAGATGGTAGAC 1851
Db 716 PheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThrPheValargLysValPro 735
QY 1852 GCTGCCGATATACAGATTCGTAGCACATTTGCGAGCTTAACTAACCCCTGCTAGAGAT 1911
Db 736 ProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAsp 755
QY 1912 CGGGAATTTTGGCTTTTACAAAAGAGAGTGGTGGTCATGTTGGTGATG 1971
Db 756 ArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeuLeu 775
QY 1972 GATTATGAGAAGTAGCTCCGCAATCAACTACATCACTTAAATTTATTTATGACATAAAC 2031
Db 776 AspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeuArgPheTyrAspIleAsp 795
QY 2032 AAGTAAATAGAGTGGATACAAAT-----TACGGC 2061
Db 796 SerValValAsnValAspAsnThrAlaValIleMetGlnGlyIleAlaSerPhePheLys 815
QY 2062 GGGCTGCAGGAATTCGGC----- 2079
Db 816 GlyLeuGlyLysValGlyGluAlaValGlyThrLeuValLeuGlyAlaAlaGlyAlaVal 835
QY 2080 -----TGTATGCTATTACTAAAAAAGATGCAACCCCAATTAATGCCAAACC--- 2127
Db 836 ValSerThrValSerGlyIleAlaSerPheLeuAsnAsnProPheGlyGlyLeuAlaIle 855

```

QY 2128 -----CAA 2130
Db 856 GlyLeuLeuValIleAlaGlyLeuValAlaAlaPhePheAlaTyrArgTyrValMetGln 875
QY 2131 TTAGAAGCGCGGAATGAGTAAAGATCAATCAATGCTAAAGCGGATGACATAGCT 2190
Db 876 IleArgSerAsnProMetLysAlaLeuTyrProIleThrThrLysAlaLeuLysAsnLys 895
QY 2191 TCACTACAAGACTATGCCAAGATTGAAGTAGTTATCATCTGCTTATAGTGAAGCTGAA 2250
Db 896 AlaLysThrSerTyrGlyGlnAsnGluGluAspGlySerAspPheAspGluAlaLys 915
QY 2251 ACAGTTAAAC-----AATAACCTTAATGCCACATTAAGACAA 2286
Db 916 LeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAlaLeuGluLysGln 935
QY 2287 CTAATAATGGCT-----AAACTAATTTAGAAATCAGCCATCAACCAAGCTAATACGGATAAA 2343
Db 936 GluLysAlaAlaIleLysLysAsnSerGlyValGlyLeuIleAlaSerAsnValSerLys 955
QY 2344 ACGACTTTTGATAATCAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACT 2403
Db 956 LeuAlaLeuArgArgGlyProLysTyr-----ThrArg 967
QY 2404 TTAGAACAACGGTCTACT 2421
Db 968 LeuGlnGlnAsnAspThr 973

RESULT 4
VGBE31
glycoprotein B - human herpesvirus 3
N:Alternate names: glycoprotein II
C:Species: human herpesvirus 3, varicella-zoster virus
C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: E27214
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: E27214
A:Molecule type: DNA
A:Residues: 1-868 <DAV>
A:Cross-references: EMBL:X04370; NID:q59989; PIDN:CAA27914.1; PID:g60020
C:Genetics:
A:Gene: 31
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:697-713/Domain: transmembrane #status predicted <TM1>
F:726-742/Domain: transmembrane #status predicted <TM2>
F:84,194,372,416,440,494,557,623,761,836,864/Binding site: carbohydrate (Asn) (covalent)

Alignment Scores:
Pred. No.: 7.39e-96 Length: 868
Score: 1777.50 Matches: 375
Percent Similarity: 61.30% Conservative: 143
Best Local Similarity: 44.38% Mismatches: 259
Query Match: 31.38% Indels: 69
Db: 1 Gaps: 16

US-09-147-052-3 (1-3261) x VGBE31 (1-868)
QY 91 AGAGAAGTTTTCGAGCGTCCAGTGTCTGAGAGAAGTCTACGTTTTATCTTTGTCCTCC 150
Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
QY 151 CCACCAAGTGGTTCACACGGTATCCGCTAGAACCCCGCGGAAGTCCCGAACCTAGA 210
Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
QY 211 AAGCCACCGAGTGGGTGAAGAAATCGGCATATTTATTAAGAGAAATATCATGTCATAT 270
Db 81 LeuGlyLysAsnPheThrGluGlyIleAlaValValTyrLysGluAsnIleAlaAlaTyr 100

```

```

271 AAATTTAAAGTACGCTTTATTTATATAAATATCATTCAGACAGCAGCATGGAGCGGAGC 330
101 LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGlySer 120
331 ACATATAGACAGATCACTAATCATGATATACAGATAGGACGCGCTTTCCATCTGAAGAGATC 390
121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140
391 ACGGATCTAATTCGCGGCAAGAGATGCTCATCTAAAGCAAGATACCTTTAGAACCAAT 450
141 ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn 160
451 GTATATGTGAAGGTTTGACAGGGATCGCGGAGAAAACAAGTACTTCTTAAACCATCA 510
161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
511 AAATTCACACGCCCGAATCTAGGCGCATGGCACAGCTAATGAGAGCTATACCGCTGG 570
181 LysTyrAsnSerValGlySerLysAlaTrpHisThrAsnAspThrTyrMetValAla 200
571 GGATCACCATGGATATATCGAACGGGAACCTCCCTCAATTCATAGTAGAGGAATGGAT 630
201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220
631 GCCGCTCTGTGTTTCGCTATTCATATTTGCAATGGCAATGCGCAGATCGCGCAACATA 690
221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSerThrGlyAspIleIleTyrMet 240
691 TCTCCATTTTATGGTCTATCCCCACAGAGGTCGCCAGAACCCATGGATATCCCCAG 750
241 SerProPhePheGlyLeuArg---AspGlyAlaTyrArgGluHisSerAsnTyrAlaMet 259
751 GATAATTTTCAAAACATAGATAGTATTTTCAATGGATTTGGACAGCGCTCGAAAAGCA 810
260 AspArgPheHisGlnPheGluGlyTyrArgGlnArgAspLeuAspThrArgAlaLeuLeu 279
811 AGCCTTCACCTCAAGCGTAACCTTCTCATCATCACATCACACTTCACAGTTGGGTGGAGCTGG 870
280 Glu---ProAlaAlaArgAsnPheLeuValThrProHisLeuThrValGlyTyrAsnTrp 298
871 GCTCCAAAATACTACTCTGTATGTTCAATGACTAAGTGGAAAGAGGTGACTGAATGTTC 930
299 LysProLysArgThrGluValCysSerLeuValLysTyrArgGluValGluAspValVal 318
931 CGTGCAACAGTTAATGGAGATACAGATTATGGCCCGTCAACTTTCCGCAACGTTTATC 990
319 ArgAspGluTyrAlaHisAsnPheArgPheThrMetLysThrLeuSerThrPheIle 338
991 AGTAATACGACTGAGTTGATCCAAATCGCATATATTAGGACAATGTATTAAACCGCAG 1050
339 SerGluThrAsnGluPheAsnLeuAsnGlnIleHisLeuSerGlnCysValLysGluGlu 358
1051 GCAGAAAGCAGCAATCGACGATATTTAGGACAAAATAATATCAGACTCAGCTCAAGTT 1110
359 AlaArgAlaIleIleAsnArgIleTyrThrArgTyrAsnSerSerHisValArgThr 378
1111 GGACATGTACAATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
379 GlyAspIleGlnThrTyrLeuAlaArgGlyGlyPheValValValPheGlnProLeuLeu 398
1171 TCCAAATCCCTGCTCATATGTACCTCAGAGATTTGATGAGAGACACACAGGACCGATGAG 1230
399 SerAsnSerLeuAlaArgLeuTyrLeuGlnGluLeuValArgGlu----- 413
1231 ATGCTCGACCTGGTAAACAATTAAGCATGCAATTTATAAGAAAATGCTACCTCATGTCA 1290
414 -----AsnThrAsnHisSer---ProGlnLysHisProThrArgAsnThr 427
1291 CGATTTGGCGCGAGATATTCGAAATGCCAACCAATAGAAAATAACATAACAGACACCACA 1350
428 ArgSerArgArgSerVal-----ProValGluLeuArgAlaAsnArg 441
1351 GCTATTAAATCGACATCGTCTGTTCAATTGCGCCATGCTCCCAATTTCTTTATGATCATATA 1410

```

```

Db      442  ThrIleThrThrSerSerValGluPheAlaMetLeuGlnPheThrThrAspHisIle 461
QY      1411 CAAACCCATTAATTAATGATGTTTATAGGATTGCCACAGCTGGTGGCAATTCAGCAAT 1470
Db      462  GlnGluHisValAsnGlnMetLeuAlaArgIleSerSerSerTrpCysGlnLeuGlnAsn 481
QY      1471 AGAGAACTGTTTATGACCAAGGATAAGATTATTCCTAGGCTACACCGAGTCCA 1530
Db      482  ArgGluArgAlaLeuTrpSerGlyLeuPheProIleAsnProSerAlaLeuAlaSerThr 501
QY      1531 ACATTAGGAAGAGAGTGCCTCAAGATGTTGGGGATGCTGCTGCTATCATGAGTGC 1590
Db      502  IleLeuAspGlnArgValylsAlaArgIleLeuGlyAspValIleSerValSerAsnCys 521
QY      1591 ACTGCTAGTAGCGGAATCC---GTCACATTGTCAAAATTCATCGGATTATCATCCTC 1647
Db      522  ProGluLeuGlySerAspThrArgIleIleLeuGlnAsnSerMetArgValSerGlySer 541
QY      1648 ACTAATACATGTTATAGCGCACCATGTTCTTATTTTATGATGGAGAAACCAAGAAAC 1707
Db      542  ThrThrArgCysTyrSerArgProLeuIleSerIleValSerLeuAsnGlySerGlyThr 561
QY      1708 ATACAGGACAACTCGGTGAAACCAACAGAGTTGCTTCCAAACGCTAGAGGCTAGAGCCA 1767
Db      562  ValGluGlyGlnLeuGlyThrAspAsnGlnLeuIleMetSerArgAspLeuLeuGluPro 581
QY      1768 TGCTCGGTAAATCATCGTAGATPATTTCCTGTTGGATCCGGTATGCTTATTTGAAAC 1827
Db      582  CysValAlaAsnHisIleArgTyrPheLeuPheGlyHisIleTyrValTyrTyrGluAsp 601
QY      1828 TATAATTTTGTAAAGTGTAGACGCTCCGATATACAGATTCTGACACATTTGTCGAG 1887
Db      602  TyrArgTyrValArgIleIleAlaValHisAspValGlyMetIleSerThrTyrValAsp 621
QY      1888 CTTAATCTAACCTGCTAGAACATCGGAAATTTTCCTTTATTCGTTTACACAAAGAA 1947
Db      622  LeuAsnLeuThrLeuLeuLysAspArgGluPheMetProLeuGlnValTyrThrArgAsp 641
QY      1948 GAGTTGCGGTGATGTTGGTATGATTATGCAAGTAGCTCGCCGCAATCAACATCAT 2007
Db      642  GluLeuArgAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnMetHis 661
QY      2008 GAACCTAAATTTATGACATAAACAAGTAATAGATGGATACAAATACCGC----- 2061
Db      662  SerLeuArgPheTyrAspIleAspLysValValGlnTyrAspSerGlyThrAlaIleMet 681
QY      2062 ---GGGCTCAGGAATTC-----GGCTGTATG----- 2076
Db      682  GlnGlyMetAlaGlnPheGlnGlyLeuGlyThrAlaGlyGlnAlaValGlyHisVal 701
QY      2077 -----GGCTGTATG-----TCPTATTACTAAAAAAGATGCA 2106
Db      702  ValLeuGlyAlaThrGlyAlaLeuLeuSerThrValHisGlyPheThrThrPheLeuSer 721
QY      2107 AACCCAAATAATGGCCAAACCAATTAGAAGCAGCGCGNATGAGTTACAGATCTAATC 2166
Db      722  AsnProPheGlyAla-----LeuAlaValGlyLeuLeuValLeuAlaGlyLeuVal 738
QY      2167 AATGCT-----AAAGCGATGACATTAGCTTCACATCAAGACTATGCCAAGATTGAAGCT 2220
Db      739  AlaAlaPhePheAlaTyrArgTyrValLeuLysLeu-LysThrSerPro---MetLysAl 757
QY      2221 AGTTTATCATCTGTTATAGTAGAAGCTGAAACAGATTAAAC-AATAACCTTAATGCAACAT 2279
Db      757  aLeuTyrProLeuThrThrLysGlyLeuLysGlnLeuProGluGlyMetAspPropheal 777
QY      2280 AGAACAACTAAATATGGCTAAACATTAATTTAGAAATCAGCCATCAACCAAGCTAATACGA 2339
Db      777  aGluLysProAsnAlaThrAspThrProIleGluIleGlyAspSerGlnAsnThrG1 797
QY      2340 TAAACGAGCTTTTGATATGAACACCCCAAT---TTAGTTGAGCATACAAAGCACTAAA 2396
;      :      :      :      :      :      :      :      :      :      :

```

```

Db      797  uProSerValAsnSerSerGlyPheAspProAspLysPheArgGluAlaGlnGluMetIleIy 817
QY      2397 A---ACCACCTTTAGAACACAGCTGCTACTAACCTTGAAGCTTTGTTCATCACTGCTTATAA 2453
Db      817  stYrMetThrLeuValSerAlaAlaGluArgGlnGlu-----Se 830
QY      2454 TCAAATTCGCATATATTTAGTGCATCTATACATAAAGCTAGTAGTTTAACTAATCACTAACAAC 2513
Db      830  rLysAlaArgLysLys-----AsnLysThrSerAlaLeuLeuThrSerAr 845
QY      2514 ACTAGATCCACTA 2526
Db      845  gleuthrGlyLeu 849

RESULT 5
VGBEPS
glycoprotein gII precursor - suid herpesvirus 1
C:Species: suid herpesvirus 1
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A29159
R:Robbins, A.K.; Dorney, D.J.; Wathen, M.W.; Whealy, M.E.; Gold, C.; Watson, R.J.; Ho
J. Virol. 61, 2691-2701, 1987
A:Title: The pseudorabies virus gII gene is closely related to the gb glycoprotein ge
A:Reference number: A29159; MUID:87284141; PMID:3039163
A:Accession: A29159
A:Molecule type: DNA
A:Residues: 1-913 <ROB>
A:Cross-references: GB:M17321; NID:g334053; PIDN:AAA47465.1; PID:g334054
A:Note: the authors translated the codon GAC for residue 860 as Asn
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-40/Domain: signal sequence #status predicted <SIG>
F:751-819/Domain: transmembrane #status predicted <TM>
F:820-913/Domain: intracellular #status predicted <INT>
F:151,261,441,516,633,697/Binding site: carbohydrate (Asn) (covalent) #status predict

Alignment Scores:
Pred. No.:      1,04e-95      Length:      913
Score:          1775.00      Matches:      373
Percent Similarity: 61.42%      Conservative: 130
Best Local Similarity: 45.54%      Mismatches: 246
Query Match:      31.33%      Indels:      70
DB:              1          Gaps:      19

US-09-147-052-3 (1-3261) x VGBEPS (1-913)
QY      52  GGTACGAACATCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAAGTGTGTTCCG 105
Db      95  GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
QY      106  AGCGTCCAGTTCTCTGAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCA 165
Db      115  AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
QY      166  ACCGTGATCCGCTAGAACCCGCGGAAATGTCCGGAACCTAGAGAAACCCACCGAGTGG 225
Db      133  ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
QY      226  GGTGAAGGAATCGCATATTTTAAAGAGAATATCAGTCCATATATAATTTAAAGTGACG 285
Db      153  ThrGluGlyIleAlaValLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
QY      286  CTTTATTATAAAATATCATTCAGACGACGACATGGACGGGACGACATATATAGACAGATC 345
Db      173  IleTyrTyrLysAsnValIleValThrValTyrPheSerGlySerThrTyrAlaAlaIle 192
QY      346  ACTAATCATATACAGATAGGACCGCGTTTCCATTCATGAAGATCATCGGATCTAATCGAC 405
Db      193  ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
QY      406  GCGAAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATCTTGNAGCG 465
Db      213  ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnAsnHisLysValThrAla 232

```


A:Reference number: JH0109; MUID:90236317; PMID:1692002

A:Accession: JH0109

A:Molecule type: DNA

A:Residues: 1-979 <GEO>

A:Cross-references: GB:M34861; NID:G330900; PIDN:AAA46086.1; PID:G330901

A>Note: glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies, box

C:Genetics:

A:Gene: gp14

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-86/Domain: signal sequence #status predicted <SIG>

F:87-979/Product: glycoprotein 14 #status predicted <NAV>

F:831-871/Region: hydrophobic

F:165,275,380,423,497,514,560,727,749,951,970/Binding site: carbohydrate (Asn) (covalent

Alignment Scores:

Pred. No.: 6,37e-95 Length: 979
Score: 1761.50 Matches: 351
Percent Similarity: 62.87% Conservative: 118
Best Local Similarity: 47.05% Mismatches: 200
Query Match: 31.09% Indels: 77
DB: 2 Gaps: 12

US-09-147-052-3 (1-3261) x JH0109 (1-979)

QY 13 AGCGGAATGCATATTTTCCTTATGTTATTTATCTATATGCT----- 54
DB 66 ArgValGlyThrIleValLeuAlaCysLeuLeuLeuPheGlySerCysValValArgAla 85
QY 55 -----ACCAACTCATCTCCGAGTACCCAAATGTGCATCA----- 90
DB 86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105
QY 91 -----AGAGAAGTGT 102
DB 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY 103 TCGAGCGTCCAGTGTCTGAGGAGAGTCTAGCTTTTATCTTGTCTCCCGCCAGGTT 162
DB 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheThrCysProProThrGly 145
QY 163 TCAACCGTGATCGCTAGAACCGCGGAAATGTCGCAACTAGAAAAGCCACCGAG 222
DB 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGGTGAGGAATCCGGATATTTAAAGAGAATATCAGTCCATATAAATTTAAAGTG 282
DB 166 PheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProThrLysPheArgAla 185
QY 283 ACGCTTTATATAAATATCATCTCAGACGACGACATGGACGGGACGACATATAGACAG 342
DB 186 AsnValTyrrLysAspIleValValThrArgValTrpLysGlyTyrrSerHisThrSer 205
QY 343 ATCACTAAATCGATATACAGTACGAGCGCGTTTCCATTTGAAGAGATCAGGATCAATC 402
DB 206 LeuSerAspArgTyrrAsnAspArgValProValSerValGluIlePheGlyLeuIle 225
QY 403 GACGCGAAGGAGATGCTCATCTATAAGCAAGATACCTTAGAACCAATATATATGTGA 462
DB 226 AspSerLysGlyLysCysSerSerLysAlaGluTyrrLeuArgAspAsnIleMethHis 245
QY 463 GCGTTTACAGGATCGGAGAGAAACAAAGTACTTCTAAACCATCAAAATCAACACG 522
DB 246 AlaTyrrHisAspAspGluAspGluValGluLeuAspLeuValProSerLysPheAlaThr 265
QY 523 CCGAATCTAGGCGTGGCAGACGACTAATGAGACG-----TATACCGTGTGGGATCA 576
DB 266 ProGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyrrValGlyTrp----Met 284
QY 577 CCATGGATATATCGAAGCGGAACTCCGTCATTTGTATAGTAGAGAAATGGATGCGCG 636
DB 285 ProTrpArgHisTyrrThrSerThrSerValAsnCysIleValGluValGluAlaArg 304

QY 637 TCTGTGTTTCCGTATTCATATTTTGAATGGCAATGGCGACATCGGACATATATCTCCA 696
DB 305 SerValTyrrProTyrrAspSerPheAlaLeuSerThrGlyAspIleValTyrrAlaSerPro 324
QY 697 TTTTATGGTCTATCCCGACAGAGCTGCGCA-----GAACCCATGGATATCCCGAG 750
DB 325 PheTyrrGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyrrAlaGln 341
QY 751 GATAATTTCAACAACATAGTAGTATTTTCAATGGATTTGGACAAAGCGTCGAAAGCA 810
DB 342 GluArgPheArgGlnValGluGlyTyrrArgProArgAspLeuAspSerLysLeuGlnAla 361
QY 811 AGCCTTCCAGTCAAGCGTAACTTTCATCATCATCATCATCATCATCATCATCATCAT 870
DB 362 GluGluProValThrLysAsnPheIleThrProHisValThrValSerTrpAsnTrp 391
QY 871 GCTCCAAAACACTACTCGTGTATGTTCAATGACTTAAGTGGAAAGAGGTGACTGAATGTTG 930
DB 382 ThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeuVal 401
QY 931 CGTGCAACAGTTAATGGGAGATACAGTTTATGGCCGCGTGAATTTTCGGCAAGCTTTATC 990
DB 402 ArgAspGluPheArgGlySerTyrrArgPheThrIleArgSerIleSerSerTyrrPheIle 421
QY 991 AGTAATACGACTGAGTTTGTCCAAATCGCATCATATATAGGCAATGTATTAAACCGGAG 1050
DB 422 SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu 441
QY 1051 GCAGAAGCAGCAATCGACGAGATATTTAGGACAAAATATAATGACAGTCACGTCAGAGTT 1110
DB 442 AlaLysGluAlaIleAspSerIleTyrrLysLysGlnTyrrGluSerThrHisValPheSer 461
QY 1111 GGACATGTACAATATTTCTGCTCTCGGGGGATTTATTGTAGCATATCAGCCTGTCTTA 1170
DB 462 GlyAspValGluTyrrTyrrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMetLeu 481
QY 1171 TCCAAATCCCTGGCTCATATGTACCTCAGAGAAATGATGAGACACACAGGACCGATGAG 1230
DB 482 SerAsnGluLeuAlaArgLeuTyrrLeuAsnGluLeuValArgSerAsnArgThrTyrrAsp 501
QY 1231 ATGCTCGACTGCTAAACAATAGCATGCAATTTATAAGAAAATGCTACTCATGTCTCA 1290
DB 502 LeuLysAsnLeuLeuAsnProAsn-----AlaAsnAsnAsnAsnAsnThrThr 517
QY 1291 CGATTCGGCGGAGATATTCGAATGACCAATAGAAAATAACATTAGAC----- 1341
DB 518 ArgArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyValHis 537
QY 1342 -----CACACCACAGCT----- 1353
DB 538 ArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAsp 557
QY 1354 -----ATTAAATCGACATCGTCTGTCTCAATTC 1380
DB 558 SerSerAsnValThrAlaLysGlnLeuGluLeuIleTyrrThrSerSerIleGluPhe 577
QY 1381 GCATGCTCCAATTTCTTATGATCATATACAAACCCATATTAATGATATGTTTATGAGG 1440
DB 578 AlaMetLeuGlnPheAlaTyrrAspHisIleGlnSerHisValAsnGluMetLeuSerArg 597
QY 1441 ATTGCCACAGCTGGTGGCGAATTTGCGAGATAGAGAACTTCTTTATGCGACAGAGGATA 1500
DB 598 IleAlaThrAlaTrpCysProLeuGlnAsnLysGluArgProLeuTrpAsnGluMetVal 617
QY 1501 AAGATTAACTCCTACGCTACAGCGAGTGCACCAATTAGGAGGAGAGTGGCTGCAAGATG 1560
DB 618 LysIleThrProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgVal 637
QY 1561 TTGGGGGATCTCGCTGTATGAGCTGCTATAGATGCGGAATCCGTCACCTTTG 1620
DB 638 LeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly---AsnValTyrrLeu 656
QY 1621 CAAAATTCATGCGAGTTATCATCCACTAAATACATGTTATAGCCGACCATTTGGTCTA 1680

```

Db      657 GlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProValThr 675
QY      1681 TTTTCATAT-----GGAGAAACCAAGAAACATACAGGCAACTCGGTGAAAC 1731
Db      676 PheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGluGluGlu 695
QY      1732 ACAGAGTTCCTCCAAAGCTGTAGAGCTGTAGAGCCATCTCGGTAAATCATCGTAGATAT 1791
Db      696 AsnGluIlePheThrGluArgLysIleGluProCysAlaLeuAsnGluLysArgTyr 715
QY      1792 TTTCTGTGTGGATCGGTTATGCTTTATTTGAAACTATAATTTGTTAAAGATGCTAGAC 1851
Db      716 PheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThrPheValArgLysValPro 735
QY      1852 GCTGCCGATATACAGATGCTAGCACATTTGTCGAGCTTAATCAACCTGCTGTAAGAT 1911
Db      736 ProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAsp 755
QY      1912 CGGGAATTTTCCTTTATCCGTTTACACAAAGAGTTCGCTGATGTTGGTATTG 1971
Db      756 ArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeuLeu 775
QY      1972 GATTATGAGAGTAGTTCGCGCAATCAACTACATGAACCTAAATTTTATGACATTAAC 2031
Db      776 AspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuArgPheTyrAspIleAsp 795
QY      2032 AAAGTAATAGAGTGGAT 2049
Db      796 SerValValAsnValAsp 801

RESULT 7
VBEOH
glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C:Species: equine herpesvirus 4
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A31880
R:Rigbio, M.P.; Cullinan, A.A.; Onions, D.E.
A:Title: Identification and nucleotide sequence of the glycoprotein gB gene of equine herpesvirus 4
A:Reference number: A31880; MUID:89125704; PMID:2915378
A:Accession: A31880
A:Molecule type: DNA
A:Residues: 1-919 <RG>
A:Cross-references: GB:M26171; NID:g341446; PIDN:AAA46106.1; PID:g514920
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence; transmembrane #status predicted <SIG>
F:29-919/Product: glycoprotein B #status predicted <GPB>
F:740-809/Domain: transmembrane #status predicted <TMN>
F:106,216,321,364,438,456,493,499,666,688/Binding site: carbohydrate (Asn) (covalent) #s

Alignment Scores:
Pred. No.: 7,23e-95 Length: 919
Score: 1760.50 Matches: 369
Percent Similarity: 58.68% Conservatives: 138
Best Local Similarity: 42.71% Mismatches: 256
Query Match: 31.08% Indels: 101
DB: 1 Gaps: 16

US-09-147-052-3 (1-3261) x VBEOH (1-919)
QY      58 AACTCATCTCCGAGTACC-----CAAAATGTC 84
Db      34 SerSerGlnProSerThrProAlaSerThrGlnSerAlaLysThrValAspGlnThrLeu 53
QY      85 ACATCAAGAGAGTCTTCGAGCGGTCCAGTGTCTGAGGAAGAGTCT-----132
Db      54 LeuProThrGluThrProAspProLeuArgLeuAlaValArgGluSerGlyIleLeuAla 73
QY      133 -----ACGTTTATCTTGTCCCCACAGTGGGTCAACCGTGTACGCTAGAA 183
Db      74 GluAspGlyAspPheTyrThrCysProProProThrGlySerThrValValArgIleGlu 93

```

```

QY      184 CCGCGCGCAAAATGTCCCAACCTAGAAAACCCAGCTGGGTGAAAGGAATCCGGATA 243
Db      94 ProProArgSerCysProLysPheAspLeuGlyArgAsnPheThrGluGlyIleAlaVal 113
QY      244 TTATTTAAAGAGATATATCAGTCCATATATAAATTTAAAGTCAGCGCTTTATTATATAAATATC 303
Db      114 IlePheLysGluAsnIleAlaProTyrLysPheArgAlaAsnValTyrTyrLysAspIle 133
QY      304 ATTCAGACGACGATACGCGGAGCAGCATATATAGACAGATCACTAATCATGATACAGAT 363
Db      134 ValValThrLysValTyrLysGlyTyrSerHisThrSerLeuSerAspArgTyrAsnAsp 153
QY      364 AGGACGCGCTTCCATTTGAAGATCACCGATCTAATCGACGCCAAGGAAGATGCTCA 423
Db      154 ArgValProValSerValGluIlePheThrLeuIleAspSerLysGlyLysCysSer 173
QY      424 TCTAAACAAAGATACCTTTAGAAACAATGTATATCTTGAAGCGTTTGACAGGGATCCGGGA 483
Db      174 SerLysAlaGluTyrLeuArgAspAsnIleMethHisAlaTyrHisAspAspGluAsp 193
QY      484 GAAACAAAGTACTTCTAAACCAATCAACACGCCCGCAATCTAGGGCATGGCAC 543
Db      194 GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaArgAlaTyrGln 213
QY      544 ACAGTAATGAGACG-----TATACCGTGTGGGATCACCATGGATATATCGAACCGGA 597
Db      214 ThrThrAsnAspThrThrSerTyrValGlyTyr---MetProTyrPheArgHisTyrThrSer 232
QY      598 ACCTCCGCTCAATGTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCTGGTATCATAT 657
Db      233 ThrSerValAsnCysIleValGluValGluAlaArgSerValTyrProTyrAspSer 252
QY      658 TTTGCAATGCCAATCGCGACATCGCGAACAATATCCATTTATGCTCTATCCCAACCA 717
Db      253 PheAlaLeuSerThrGlyAspIleValTyrThrSerProPheTyrGlyLeuArg---Ser 271
QY      718 GAGGCTGCCGACGACCCATGGGATATCCCGAGTAATTTCAACAACTAGATAGCTAT 777
Db      272 AlaAlaGlnLeuGluHisAsnSerTyrAlaGlnGluArgPheArgGlnValGluGlyTyr 291
QY      778 TTTTCAATGGATTTGGACAAGCGCTCGAAGCAAGCGCTTCCAGTCAAGCGTAATCTTCTC 837
Db      292 GlnProArgAspLeuAspSerLysLeuGlnAlaGlyProValThrLysAsnPheIle 311
QY      838 ATCACAATCACATTCACAGTTGGTGGAGTGGCTCCAAAACACTACTCGTGTATGTTCA 897
Db      312 ThrThrProHisValThrValSerTyrPheAsnThrThrGluLysLysIleGluAlaCysThr 331
QY      898 ATGACTAAGTGGAAAGAGGTAGTGAATGTTGCGTGCAACAGTAAATGGAGATACAGA 957
Db      332 LeuThrLysTyrLysGluValAspGluLeuValArgAspLysPheArgGlySerTyrArg 351
QY      958 TTTATGCGCGCTCACTTTCCGCAAGCTTATCAGTAATACGACTGAGTTTTCATCCAAAT 1017
Db      352 PheThrIleArgSerIleSerThrPheIleSerAsnThrThrGlnPheLysLeuGlu 371
QY      1018 CGCATCATATTAGGCAATGTATTAACCGGAGGAGGACGATCCAGCATATTT 1077
Db      372 AspAlaProLeuThrAspCysValSerLysGluAlaLysAspAlaIleAspSerIleTyr 391
QY      1078 AGCACAATAATATGACAGTCCAGTCCAGTGGAGATGTACAATATTTCTTGGCTCTC 1137
Db      392 ArgLysGlnTyrGluSerThrHisValPheSerGlyAspValGluPheTyrLeuAlaArg 411
QY      1138 GGGGATTTTATGTCAGCATATCAGCTGTCTTATCCAAATCCCTGGCTCATATGATCC 1197
Db      412 GlyGlyPheLeuIleAlaPheArgPrometIleSerAsnGluLeuAlaArgLeuTyrLeu 431
QY      1198 AGAGAAATGATGAGACACACAGGACCGATGAGATGCTCGACCTGGTAAACAAATAGAT 1257
Db      432 AsnGluLeuValArgSerAsnArgThrTyrAspLeuLysAsnLeuLeu---AsnProAsn 450

```

```
QY 1258 GCAATTTATAAGAAAATGCTACCTCA----- 1284
Db 451 AlaAsnHisAsnThrAsnArgThrArgSerLeuLeuSerIleProGluProThrPro 470
QY 1285 -----TTGCACGATTGCGCGAGATATTCGAAAT 1314
Db 471 ThrGlnGluSerLeuHisArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGlu 490
QY 1315 GCACCAAT-----AGAAATATACATTAGACGACACACAGCTATTAAATCG 1362
Db 491 AlaAlaAsnSerThrAsnSerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThr 510
QY 1363 ACATCGCTGTTCAATTCGCCATCTCCCAATTTCTTATGATCATATACAAACCATAT 1422
Db 511 ThrSerIleGluPheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisVal 530
QY 1423 AATGATATGTTAGTAGGATGCCACAGCTTGGTGGCAATTCGCAATAGAGAACTGTT 1482
Db 531 AsnGluMetLeuSerArgIleAlaThrAlaTyrCysThrLeuGlnAsnLysGluArgThr 550
QY 1483 TTATGGCAGCAGGATTAAGATTAATCTTAGCGGTACAGCGATGCAACATTAGGAAG 1542
Db 551 LeuTrpAsnGluMetValLysValAsnProSerAlaIleValSerAlaThrLeuAspGlu 570
QY 1543 AGAGTGGCTGCAAGAGATCTGGGGATCTCGCTGCTATCGAGCTGCACTCTATAGAT 1602
Db 571 ArgValAlaAlaArgValLeuGluAspValIleAlaIleThrHisCysValLysIleGlu 590
QY 1603 GCGGAATCCGTCACCTTTGCAAAATCTATGCGGATTATCATCATCCACATTAATCATGTTAT 1662
Db 591 Gly---AsnValTyrLeuGlnAsnSerMetArg---SerSerAspSerAsnThrCysTyr 608
QY 1663 AGCCGACCATGGTGTCTATTTTCATATATGAGAAAC-----CAAGAAACATACAG 1713
Db 609 SerArgProProValThrPheThrIleThrLysAsnAlaAsnSerArgGlyThrIleGlu 628
QY 1714 GCACAACTCGGTGAAAAACAGAGTGTCTTCCACGCTAGAGGCTGAGAGCCATGCTCG 1773
Db 629 GlnGlnLeuGluGluAsnGluValTyrThrGluArgLysLeuIleGluProCysAla 648
QY 1774 GCTAATCATCTAGATATTTCTGTTGGATCCGGTTATGCTTTATTTGAAACATATAAT 1833
Db 649 IleAsnGlnLysArgTyrPheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThr 668
QY 1834 TTTGTTAAGGTAGTAGCGTCCGATATACAGATTGTAGACATTTGTCGAGCTTAAAT 1893
Db 669 TyrValArgLysValProThrGluIleGluValIleSerThrTyrValGluLeuAsn 688
QY 1894 CTAACCCCTGCTAGAGATCGGAAATTTTGGCTTTATCCGTTTACACAAACAGAGTTG 1953
Db 689 LeuThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeu 708
QY 1954 CGTGATGTTGGTATTGGATTATGAGAAGTAGCTCGCCGAATCAACTACATCAAACTT 2013
Db 709 GluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeu 728
QY 2014 AAATTTTATGATTAACAAGTATAGAGTGGAT----- 2049
Db 729 ArgPheTyrAspIleAspSerValValAsnValAspAsnThrAlaValIleMetGlnGly 748
QY 2050 -----ACAAATACCGCGGGCTCGAGGAATTCGGCTGTATGCTATTACTTAAAAAAGAT 2103
Db 749 IleAlaThrPhePheGlyLeuGlyLysValGly----- 760
QY 2104 GCMAACCCAAATATAGGCCAAACCAATAGAACGACGCGGAATGGAGTTAACAGATCTA 2163
Db 761 -----GluAlaValGlyThrLeuValLeuGlyAla 770
QY 2164 ATCAATGCTAAGCGATGACATAGCTTCTACTACAGACTATGCCAAG----- 2211
Db 771 AlaGlyAlaValSerThrValSerGlyIleAlaSerPheIleAsnAsnProPheGly 790
QY 2212 -----ATTGAACCTAGTTTATCATCTGCTTATATAGTGAACCTGAA 2250
```

```
Db 791 GlyLeuAlaIleGlyLeuLeuValIleAlaGlyLeuValAlaAlaPhePheAlaTyrArg 810
QY 2251 ACAGTT-----AACAAATACCTTTATGTCACCACTTAGAACAACATAAAATGGCT 2298
Db 811 TyrValMetGlnLeuArgSerAsnProMetLysAlaLeuTyrProIleThrThrArgSer 830
QY 2299 AAACTAATTTAGATATCAGCATCAACCAAGCTAATACGGATAAAACGACTTTTATAT 2358
Db 831 LeuLysAsnLysAlaLysAlaSerTyrGlyGlnAsnAspAspAspThrSerAspPhe 850
QY 2359 GAACACCCCAATTTAGTTGAAGCA-----TACAAAGCACTAAAAACCACT 2403
Db 851 AspGluAlaLysLeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAla 870
QY 2404 TTAGAACCAACGT 2415
Db 871 LeuGluLysGln 874

RESULT 8
T42576
probable envelope protein 33 - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42576
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: 222173; MUID:98264497; PMID:9603335
A:Accession: T42576
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-975 <TEL>
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59549.1; PID:g2605977
A:Experimental source: strain NS80567
C:Genetics:
A:Gene: 33
C:Superfamily: herpesvirus glycoprotein B

Alignment Scores:
Pred. No.: 1,42e-94 Length: 975
Score: 1755.50 Matches: 368
Percent Similarity: 58.58% Conservative: 138
Best Local Similarity: 42.59% Mismatches: 257
Query Match: 30.99% Indels: 101
DB: 2 Gaps: 16

US-09-147-052-3 (1-3261) x T42576 (1-975)
QY 58 AACTCATCTCCGAGTACC-----CAAAATCTG 84
Db 90 SerSerGlnProSerThrProAlaSerThrGlnSerAlaLysThrValAspGlnThrLeu 109
QY 85 ACATCAAGAGAAGTTCCTCGAGCGTCCAGTGTCTGAGGAAGAGTCT----- 132
Db 110 LeuProThrGluThrProAspProLeuArgLeuAlaValArgGluSerGlyIleLeuAla 129
QY 133 -----ACGTTTATCTTTGTCCTCCACCGTGGGTTCACCGTGTACCGTCTAGAA 183
Db 130 GluAspGlyAspPheTyrThrCysProProThrGlySerThrValValArgIleGlu 149
QY 184 CGCCGCGAAAATGTCGCGAACCTAGAAAAGCCACCGAGTGGGTGGGAGGAATCGGATA 243
Db 150 ProProAspSerCysProLysPheAspLeuGlyArgAsnPheThrGluGlyIleAlaVal 169
QY 244 TTATTTAAGAGATATCAGTCCATATAATTTAAAGTGACCGCTTTATTATATAAATATC 303
Db 170 IlePheLysGluAsnIleAlaProTyrLysPheArgAlaAsnValTyrTyrLysAspIle 189
QY 304 ATTCAGACGACGACATGACGGGACGACATATAGACATCAGATCAGTCAATCGATATACAGT 363
Db 190 ValValThrLysValTrpLysGlyTyrSerHisThrSerLeuSerAspArgTyrAsnAsp 209
```

QY 364 AGGACGCCGCTTCCATTGAGAGATACCGGATCTAATFCGACGCCAAAGGAAGATGCTCA 423
DB 210 ArgValProValSerValGluGluLeuPheThrLeuIleAspSerLysGlyLysCysSer 229
QY 424 TCTAAAGCAAGTACCTTTAGAAAACAATCTATATGTTGAAGCGTTTGACAGGGATCGCGGA 483
DB 230 SerLysAlaGluTyrLeuArgAspAsnIleMetHisAlaTyrHisAspAspGluAsp 249
QY 484 GAAAAACAAGTACTTCTAAACCACTAAAATTTCAACACGCCGCCGAATCTAGGGCATGCAC 543
DB 250 GluValGluLeuAspLeuValProSerLysPheAlaThrProGlyAlaArgAlaTyrPheGln 269
QY 544 ACAGTAAATGAGACG-----TATACCGTGTGGGGATCACCATGGATATATACGACGGGA 597
DB 270 ThrThrAsnAspThrThrSerTyrValGlyTyr---MetProTyrAspHisTyrThrSer 288
QY 598 ACCTCCGCTCAATGTATAGTAGAGAAATGGATGCCGCTCTGTTTCGTTATTCATAT 657
DB 289 ThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyrAspSer 308
QY 658 TTGCAATGGCAATGGCGACATCGCGAACAATATCTCCATTTTATGTCCTATCCGCCACCA 717
DB 309 PheAlaLeuSerThrGlyAspIleValTyrThrSerProPheTyrGlyLeuArg---Ser 327
QY 718 GAGGCTGCCGACAGAACCCATGGGATATCCCGAGGATAATTTCAACAACTAGATAGCTAT 777
DB 328 AlaAlaGlnLeuGluHisAsnSerTyrAlaGlnGluArgPheArgGlnValGluGlyTyr 347
QY 778 TTTTCAATGATTTGGACAAGCGTCGAAAGCGCTTCCAGTCAAGCGCTCAACTTCTC 837
DB 348 GlnProArgAspLeuAspSerLysLeuGlnAlaGlyGluProValThrLysAsnPheIle 367
QY 838 ATCATCATCATCTCACATTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 897
DB 368 ThrThrProHisValThrValSerThrPheSerThrGluLysLysIleGluAlaCysThr 387
QY 898 ATGACTAAGTGAAGAGGTGACTGAATGTTCCGTGCAACAGTAAATGGGAGATACAGA 957
DB 388 LeuThrLysTrpLysGluValAspGluLeuValArgAspGluPheArgGlySerTyrArg 407
QY 958 TTTATGGCCGCTGAACTTTCCGCAACGTTTATCAGTAATACGACTGAGTTGATCCAAAT 1017
DB 408 PheThrIleArgSerIleSerSerThrPheIleSerAsnThrThrGlnPheLysLeuGlu 427
QY 1018 CGCATCATATAGCAATATGTTAAACCGGACAGAGACGACATCCAGCAGATATTT 1077
DB 428 AspAlaProLeuThrAspCysValSerLysGluAlaLysAspAlaIleAspSerIleTyr 447
QY 1078 AGGACAAAATATAATGACAGTCACGTCACAGTTGGACATGTACAATATTTCTTGGCTCTC 1137
DB 448 ArgLysGluTyrGluSerThrHisValPheSerGlyAspValGluPheTyrLeuAlaArg 467
QY 1138 GGGGATTTATGTACCATATCAGCTGTCTCTATCCAAATCCCTGGCTCATATGACCTTC 1197
DB 468 GlyGlyPheLeuIleAlaPheArgProMetIleSerAsnGluLeuAlaArgLeuTyrLeu 487
QY 1198 AGAGATTGATGACAGACAAACAGGACCGATGATGCTCGAGCTGTAACAACTAAGCAT 1257
DB 488 AsnGluLeuValArgSerAsnArgThrTyrAspLeuLysAsnLeuLeu---AsnProAsn 506
QY 1258 GCAATTTATACAAAATGCTACCTCA----- 1284
DB 507 AlaAsnHisAsnThrAsnArgThrArgArgSerLeuLeuSerIleProGluProThrPro 526
QY 1285 -----TTGTCACGATTCGCGCGGAGATATTCGAAAT 1314
DB 527 ThrGlnGluSerLeuHisArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGlu 546
QY 1315 GCACCAAT-----AGAAAAATAACATTAGACGACACACAGCTATTAAATCG 1362
DB 547 AlaAlaAsnSerThrAsnSerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThr 566

QY 1363 ACATCGCTGTTCATTCGCATCTCCAAATTTCTTTATGATCATATACAAACCCATATT 1422
DB 567 AlaSerSerIleGluPheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisVal 586
QY 1423 AATGATATGTTAGTAGATTGGCACAGCTTGTGTGCGAATTTGCAGATACAGACTTGTT 1482
DB 587 AsnGluMetLeuSerArgIleAlaThrAlaTyrCysThrLeuGlnAsnLysGluArgThr 606
QY 1483 TTATGGCAGCAAGGATTAAGATTAATCTAGCGCTACAGCGAGTGCAGACTTAGGAAGG 1542
DB 607 LeuTrpAsnGluMetValLysValAsnProSerAlaIleValSerAlaThrLeuAspGlu 626
QY 1543 AGAGTGGCTCAAGATGTTGGGGATCTCGCTGCTGTATCGAGCTGCAGCTGTATAGAT 1602
DB 627 ArgValAlaAlaArgValLeuGlyAspValIleAlaIleThrHisCysValLysIleGlu 646
QY 1603 GCGGAATCCGCTACTTTGCAAAATTTATCGAGTATTACATCCACTCAATACATCTTAT 1662
DB 647 Gly---AsnValTyrLeuGlnAsnSerMetArg---SerSerAspSerAsnThrCysTyr 664
QY 1663 AGCGGACCATTTGTTCTATTTTCTATATGGGAGAAAC-----CAAGAAACATACAG 1713
DB 665 SerArgProProValThrPheThrIleThrLysAsnAlaAsnSerArgGlyThrIleGlu 684
QY 1714 GGAACTCGTGGTGAACCAACAGAGTTGCTTCCAACTAGAGCTGTAGAGCATCTCG 1773
DB 685 GlyGlnLeuGlyGluGlnAsnGluValTyrThrGluArgLysLeuIleGluProCysAla 704
QY 1774 GCTAATCATCGTAGATATTTCTGTTGGTATCGGTTATGCTTTATTTGAAACTATAAT 1833
DB 705 IleAsnGlnLysArgTyrPheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThr 724
QY 1834 TTTGTAAGATGATGACGCTGCCGATATACAGATGTGTAGCAGACTTTCTCGAGCTTAT 1893
DB 725 TyrValArgLysValProProThrGluIleGluValIleSerThrTyrValGluLeuAsn 744
QY 1894 CTAACCTCTAGAGATCGGAAATTTTCCTTTATCGTTTACACAAAAGAGTGTG 1953
DB 745 LeuThrLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeu 764
QY 1954 CGTGATTTGGTATTTGATTTATCGAGAGTGTGCTCGCGCACTACACTACACTT 2013
DB 765 GluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisAlaLeu 784
QY 2014 AAATTTTATGACATAAACAAGTAAATAGAGTGGAT----- 2049
DB 785 ArgPheTyrAspIleAspSerValValAsnValAspAsnThrAlaValIleMetGlnGly 804
QY 2050 -----ACAATTCGCGGGCTCGAGGATTCGCTGTATGCTTATTACTAAAAAGAT 2103
DB 805 IleAlaThrPhePheLysGlyLeuGlyLysValGly----- 816
QY 2104 GCAACCCCAATTAATGGCAACCACTTAGAGCAGCGCGATGGAGTATACAGATCTA 2163
DB 817 -----GluAlaValGlyThrLeuValLeuGlyAla 826
QY 2164 ATCAATGCTAAGCGATGACATTAAGTCTACTACAAGACTATGCCAAG----- 2211
DB 827 AlaGlyAlaValValSerThrValSerGlyIleAlaSerPheIleAsnAsnPropheGly 846
QY 2212 -----ATTGAAGTATTATCATCTGCTTATAGTGAAGCTGAA 2250
DB 847 GlyLeuAlaIleGlyLeuLeuValIleAlaGlyLeuValAlaAlaPheAlaTyrArg 866
QY 2251 ACAGTT-----AACAATAACCTTAATCAACATTAAGAACACTAAAAATGGCT 2298
DB 867 TyrValMetGlnLeuArgSerAsnProMetLysAlaLeuTyrProIleThrThrArgSer 886
QY 2299 AAACTATTATTAGAAATCAGCCATCAACCAAGCTAATACGGATAAAAGCACTTTTGAAT 2358
DB 887 LeuLysAsnLysAlaLysAlaSerTyrGlyGlnAsnAspAspAspThrSerAspPhe 906
QY 2359 GAACACCAAAATTTAGTTGAAGCA-----TCAAAAGCACTAAAAACCACT 2403

Db 907 AspGluAlaLysLeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAla 926
Qy 2404 TTAGAACACAGCT 2415
Db 927 LeuGluLysGln 930

RESULT 9

348474
glycoprotein B - feline herpesvirus 1 (fragment)
C:Species: feline herpesvirus 1
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C:Accession: B48474
R:Spatz, S.J.; Maes, R.K.
Virology 197, 125-136, 1993
A:Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B and a
A:Reference number: A48474; MUID:94025559; PMID:8212548
A:Contents: C-27
A:Accession: B48474
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-943 <SPA>
A:Cross-references: GB:S66371; NID:G435792; PIDN:AAB28559.1; PID:G435794
A:Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBI:P:138807)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Alignment Scores:
Pred. No.: 7.54e-94 Length: 943
Score: 1743.00 Matches: 377
Percent Similarity: 57.54% Conservative: 161
Best Local Similarity: 40.32% Mismatches: 267
Query Match: 30.77% Indels: 130
DB: 2 Gaps: 21
US-09-147-052-3 (1-3261) x B48474 (1-943)

Qy 7 TATTTTAGCGGAATTGCATATT----- 30
Db 22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
Qy 31 -----TTCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
Qy 43 ATCTATAT-----GGTACCACTCATCTCCG 69
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
Qy 70 AGTACCCAAATGTGACATCAAGAGAAGTT----- 99
Db 82 ProArgArgThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
Qy 100 -----GTTTCGAGCGTCCAGTTGTCGAGGAAGAG----- 129
Db 102 ProProIleCysArgTyrGluGluAlaAlaArgAlaSerGlnIleGluAlaAsnGlyPro 121
Qy 130 TCTACGTTTATCTTTGTCCTCCACCACTGCGGTTCACCGGTGATCCGCTAGAACCGCGC 189
Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
Qy 190 CGAAATGTCCTCCGACCTAGAAAGCCACCGAGTGGGTGAAGGAATCGCATATTATT 249
Db 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
Qy 250 AAAGAGATATCAGTCCATATAAATTTAAAGTGACCGCTTTATTATATAAATATCATTCAG 309
Db 162 LysGlnAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
Qy 310 ACGACGACATGACGGGACGACATATAGACAGATCACTAATCGATATACAGATAGGACG 369
Db 182 ThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201

Qy 370 CCCGTTTCCATTGAAGAGATCACGGATCTAATCCAGCGGCAAAAGGAAGATGCTCATCTAAA 429
Db 202 ProValLysGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSerLys 221
Qy 430 GCAAGATACCTTAGAAAACATGTATGTTCAACGGTTTGAACGGATCGCGGAGAAAA 489
Db 222 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgGluAspProArg 241
Qy 490 CAAGTACTTCTPAAACACCAATCAACACCGCCGAATCTAGGCGATCGCACACGACT 549
Db 242 GluLeuProLeuLysProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260
Qy 550 AATCAGACGTATACCGTGTGGGATCACCATGGATATATCAACGGGACCGTCCGTCAT 609
Db 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn 280
Qy 610 TGTATAGTAGAAGAAATGGATGCCGCTCTGTTTCGTTATTCATATTTTTCATGCGC 669
Db 281 CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300
Qy 670 AATGGCGACATCGCGAACATATCTCCATTTTATGGTCTATCCCGACACGAGCGTCCGCA 729
Db 301 ThrGlyAspValIleHisMetSerProPhePheGlyLeuArg---AspGlyAlaHisVal 319
Qy 730 GAACCCATGGATATCCCGAGGATAATTTCAACACATAGATAGTATTTTCAATGGAT 789
Db 320 GluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrProIleAsp 339
Qy 790 TTGCACAAGCGTCGAAAGCAAGCTTCCAGTCAAGCGTAACCTTCATCATCACATCACAC 849
Db 340 LeuAspThrAspTyrThrGlyAla---ProValSerArgAsnPheLeuGluThrProHis 358
Qy 850 TTCACAGTTGGTGGGACTGGCGCTCCAAAATCTACTGCTGATGTTCAATGACTAAGTGG 909
Db 359 ValThrValAlaTrpAsnTrpThrProLysSerGlyArgValCysThrLeuAlaLysTrp 378
Qy 910 AAAGAGTGTAGTAATGTCGCTGCGCAACAGTTAATGGGAGATACAGATTTATGCGCGCT 969
Db 379 ArgGluIleAspGluMetLeuProMetAsnIle---GlySerTyrArgPheThrAlaLys 397
Qy 970 GAACCTTCGGCAACGTTTATCAGTAATACGACTGAGTTTGATCAATCGCATCATATTA 1029
Db 398 ThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeu 417
Qy 1030 GGACAAATGTATTAACCGGAGGAGACAGACAGATCGAGCAGATATTTAGGCAAAATAT 1089
Db 418 GlyAspCysAlaThrLysGluAlaAlaIleAspArgIleTyrLysSerLysTyr 437
Qy 1090 AATCAGACTCAGCTCAAGTTGGACATGTACAATATTCTTGGCTCTCGGGGGATTTATT 1149
Db 438 SerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeu 457
Qy 1150 GTACATATCAGCGCTTCTATCAAAATCCCTGGCTCATATGTAACCTCAGAGAAATGATG 1209
Db 458 IleAlaPheArgProMetIleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeuAla 477
Qy 1210 AGACACACAGGACGATGAGATGCTCGCATGTTAAACAATAAGCATGCAATTTATAAG 1269
Db 478 ArgSerAsnArgThr-----ValValAspLeu-----SerAlaLeuLeuAsn 491
Qy 1270 AAAAATGCTACTCATTTGTCAGGATCGCGGAGATATTCGAAATGCAAAATAGAAAA 1329
Db 492 ProSerGlyGluThrValGlnArgThrArgSerValProSerAsnGlnHisArg 511
Qy 1330 ATA-----ACATTAGACGACACACGACTATTAAA 1359
Db 512 SerArgArgSerThrIleGluGlyIleGluThrValAsnAsnAlaSerLeuLeuLys 531
Qy 1360 TCGACATCGCTGTTTCAATTCGCCATGCTCCAAATTTCTTATGATCATATCAAAACCCAT 1419
Db 532 ThrThrSerSerValGluPheAlaMetLeuGlnPheAlaTyrAspTyrIleGlnAlaHis 551
Qy 1420 ATTAATGATATGTTTAGTAGGATGCCACAGCTTGGTCGGAATGTCAGAAATAGAGAACTT 1479

Db 552 ValAsnGluMetLeuSerArgIleAlaThrAlaTyrCysThrLeuGlnAsnArgGluHis 571
QY 1480 GTTTTATGGCAGGAGGATAAGATTAACTAGCGCTACAGCGAGTGCACACATTAGGA 1539
Db 572 ValLeuTprThrGluThrLeuLysLeuAsnProGlyGlyValValSerMetAlaLeuGlu 591
QY 1540 AGGAGAGTGGCTGCAAGAGATGTTGGGGATGCTGCTGCTATCGAGCTGCACCTATATA 1599
Db 592 ArgArgValSerAlaArgLeuGluGlyAspAlaValAlaValThrGlnCysValAsnIle 611
QY 1600 GATCGGAATCCGCTACCTTTGCAAAATCTATCGAGTTCATCATCCATCATACATGT 1659
Db 612 SerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrThrCys 631
QY 1660 TATAGCCGACATGTTGTTCTATTTTCATATCGAGAAACCAAGAAACATACAGGACAA 1719
Db 632 TyrSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluTyrIleGluGlyGln 651
QY 1720 CTCGGTGAACCAACAGAGTTCCTCCACGCTAGAGGCTGCTAGAGCATGCTCGGCTAAT 1779
Db 652 LeuGlyGlnAsnAsnGluLeuLeuValGluArgLysLeuIleGluProCysThrValAsn 671
QY 1780 CATCGTAGATATTTCTGTTGGATCCGGTTATGCTTTATTTGAAACTATAATTTGTT 1839
Db 672 AsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyrVal 691
QY 1840 AAGATGCTAGACCTCCCGATATACAGATTCGTAGACATTTCTCGAGCTTAATCAACC 1899
Db 692 ArgLysValProLeuSerGluIleGluLeuIleSerAlaTyrVal---IleLysSerThr 710
QY 1900 CTGCTAGAGATCGGAAATTTTGCCTTTATCCGTTTACACAAAGAGTTCGCTGAT 1959
Db 711 LeuLeuGluAspArgGluPheLeuHis---SerSerTyrThrArgAlaGluLeuGluAsp 729
QY 1960 GTTGGTGTATTGATTATGATGAGAGTATGCTCGCGCAATCAACTACATCAACTTAATTT 2019
Db 730 ThrGlyProPheAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuLysPhe 749
QY 2020 TATGACATAAACAATAAGATGAGTGGATACAAAT-----2055
Db 750 TyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMetAla 769
QY 2056 -----TACCGGGCTGCAGGAATTCGGCTGT-----2082
Db 770 AsnPheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGlyAla 789
QY 2083 -----ATGCTATTACTAAAGATGCAACCCCAATAATGCCAAACC 2127
Db 790 AlaSerAlaValIleSerThrValSerGlyValSerPheLeuAsnAsnProPheGly 809
QY 2128 CAATTAGAACGACCGGATGAGTAAACAGATCTAATCAATGCT-----2172
Db 810 AlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleValAlaAlaPheLeuAlaTyrArg 829
QY 2173 -----AAAGCGATG-----ACATTAGCTTCA 2193
Db 830 TyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArgAsn 849
QY 2194 CTACAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTAGAGCTGAACA 2253
Db 850 LeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyLysAspSerAspProGlyVal 869
QY 2254 GTTAACAATAACCTTAATGCAACATTAAGCAACTAAAAATGGCTAAA---ACTAATTTA 2310
Db 870 AspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSerLeu 889
QY 2311 GAATCAGCCATCAACCAAGCTAATACGGAATAAAGCACTTTTGATAATGAACACCCAAAT 2370
Db 890 ValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyProAla 908
QY 2371 TTAGTTGAAGCATAC-----AAAGCACTAAACACC-----ACTTTAGAACAACGT 2415
Db 2415

Db 909 IleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgGlyProLysTyrGlnArg 928
QY 2416 GCTACTAACCTTGAAGTTTGTCTCATCACTGCTTATAATCAAAATT 2460
Db 929 LeuAsnAsnLeuAspSerGlyAspAspThrGluThrAsnLeuVal 943
RESULT 10
VGBE2H
glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)
C;Species: equine herpesvirus 1
C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
C;Accession: A31241; JQ0008
R;Whalley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Bell, C.W.; Woodworth, L
J. Gen. Virol. 70, 383-394, 1989
A;Title: Identification and nucleotide sequence of a gene in equine herpesvirus 1
A;Reference number: A31241; MUID:89279217; PMID:2543744
A;Accession: A31241
A;Molecule type: DNA
A;Residues: 1-980 <WHA>
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-85/Domain: signal sequence #status predicted <SIG>
F;70-78/Domain: transmembrane #status predicted <TN1>
F;86-980/Product: glycoprotein B #status predicted <NAT>
F;853-868/Domain: transmembrane #status predicted <TN2>
F;165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (C)
Alignment Scores:
Pred. No.: 9,25e-94 Length: 980
Score: 1741.50 Matches: 379
Percent Similarity: 55.56% Conservative: 136
Best Local Similarity: 40.88% Mismatches: 269
Query Match: 30.74% Indels: 143
DB: 20 Gaps: 1
US-09-147-052-3 (1-3261) x VGBE2H (1-980)
QY 13 AGCGGAATTCGATATTTTCTTATAGTTATTTCTATATGTT-----54
Db 66 ArgValGlyThrIleValLeuAlaCysLeuLeuPheGlySerCysValValArgAla 85
QY 55 -----ACGACTCATCTCCGAGTACCCAAATGTGACATCA-----90
Db 86 ValProThrThrProSerProProThrSerThrProThrSerMetSerThrHisSerHis 105
QY 91 -----AGAGAAGTTGTT 102
Db 106 GlyThrValAspProThrLeuLeuProThrGluThrProAspProLeuArgLeuAlaVal 125
QY 103 TCGAGCGTCCAGTTGCTGAGGAAGAGTCTACGTTTATCTTTTCCCGCCACCATGGGT 162
Db 126 ArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysProProThrGly 145
QY 163 TCAACCGTGATCCGCTAGAACCCCGCGAATAATGTCCCGAACCTAGAAAGCCACCGAG 222
Db 146 SerThrValValArgIleGluProProArgThrCysProLysPheAspLeuGlyArgAsn 165
QY 223 TGGGCTGAAGGAATCGGATATTTAAAGAGAAATATCATGCTCCATATAATTTAAAGTG 282
Db 166 PheThrGluGlyIleAlaValIlePheLysGlnAsnIleAlaProTyrLysPheArgAla 185
QY 283 ACGCTTTATATAAATATCATTTCAGACGACATCGGACGACGACATATAGACAG 342
Db 186 AsnValTyrTyrLysAspIleValThrArgValThrLysGlyTyrSerHisThrSer 205
QY 343 ATCACTAATCGATATACAGATGAGCCCGCTTCCATTGAAGAGATCAGGATCTAATC 402
Db 206 LeuSerAspArgTyrAsnAspArgValProValSerValGluGluIlePheGlyLeu 225
QY 403 GACGGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAGAACATATGATGTTGAA 462
Db 226 AspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsnIleMetHis 245

```
QY 463 GCGTTTGCAGGATCGGGAGAAAAACAAGTA--CTTCTAAACCATCAAAATTCAAC 519
Db 246 AlaTyrHisAspAspGluValGluLeuAspLeuCysArgProSer-----LeuGln 264
QY 520 ACGCCGAATCTAGGCATGGCCACAGACTAATGAGACG-----TATACCGTGTGGGA 573
Db 265 LeuArgGlyAlaArgAlaIrpGlnThrThrAsnAspThrThrSerTyrValGlyTyr--- 283
QY 574 TCACATGGATATATCGAAGCGGAACCTCGCTCAATGTATATAGAGAAATGGATGCC 633
Db 284 MetProTArgHisTyrThrSerThrSerValAsnCysIleValGluGluValGluAla 303
QY 634 CGCTCTGTGTTCCGTATTTCATATTGCAATGCGCAATGGCGACATCGCAACATATCT 693
Db 304 ArgSerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrAlaSer 323
QY 694 CCATTTTATGGTATATCCACCAGAGGCTCGCGCA-----CAACCCCATGGATATCCC 747
Db 324 ProPheTyrGlyLeu-----ArgAlaAlaAlaArgIleGluHisAsnSerTyrAla 340
QY 748 CAGGATAATTTCAACAACATAGATAGCTATATTTTCAATGGATTTGGCAAGCCGTCGAAA 807
Db 341 GlnGluArgPheArgGlnValGluGlyTyrArgProArgAspLeuAspSerLysLeuGln 360
QY 808 GCAAGCCTTCCAGTCAAGCGTAACTTCTCATCATCATCACATTCACAGTTGGGTGGGAC 867
Db 361 AlaGluGluProValThrLysAsnPheIleThrThrProHisValThrValSerTyrPasn 380
QY 868 TGGGCTCCAAAATACTACTCGTGTATGTTCAATGACTAAGTGGAAAGAGGTGACTGAAATG 927
Db 381 TrpThrGluLysLysValGluAlaCysThrLeuThrLysTrpLysGluValAspGluLeu 400
QY 928 TTCGGTGCACAGTTAATGGGAGATACAGATTTATGCCCGTGAATTTTCGGCAAGCTTT 987
Db 401 ValArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerSerThrPhe 420
QY 988 ATCAGTAATACGACTGAGTTGATCCAAATCGCATCATATATAGGCAATGTATTAACGC 1047
Db 421 IleSerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLys 440
QY 1048 GAGCGAGAAGCAGCAATCGAGCAGATATTTAGACAAATATATATGACAGTCAAGTCAAG 1107
Db 441 GluAlaLysGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPhe 460
QY 1108 GTTGGCAGTCACAATATTTCTTGGCTCGCGGGATTTATTTGATAGATATCAGCTGTT 1167
Db 461 SerGlyAspValGluTyrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMet 480
QY 1168 CTATCCAAATCCCTGGCTCATATGTACCTCAGAGAATTGATGAGACACAGACCGCAT 1227
Db 481 LeuSerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyr 500
QY 1228 GAGATGCTCGACTGGTAAACAATAAGCATGCAATTTATAGAAAATGCTACTCATTTG 1287
Db 501 AspLeuLysAsnLeuLeuAsnProAsn-----AlaAsnAsnAsnAsnThr 516
QY 1288 TCACGATTCGGGGCAGATATTCGAATGACCAAAATAGAAAATACATTAGAC----- 1341
Db 517 ThrArgArgArgSerLeuLeuSerValProGluProGlnProThrGlnAspGlyVal 536
QY 1342 -----GACACCCACAGCT----- 1353
Db 537 HisArgGluGlnIleLeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThr 556
QY 1354 -----ATTAATTCGACTCGTCTGTTCAA 1377
Db 557 AspSerSerAsnValThrAlaLysGlnLeuGluLeuIleLysThrThrSerSerIleGlu 576
QY 1378 TTCGCCATGCTCCAATTTCTTTATGATCATATACAAACCCCATATTAATGATGTGTTAGT 1437
Db 577 PheAlaMetLeuGlnPheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSer 596
QY 1438 AGGATTTGCCACAGCTTGGTCCGAATTCAGATAGAGAACTGTTTATGGCAGGAGG 1497
Db 597 ArgIleAlaThrAlaTyrCysThrLeuGlnAsnLysGluArgThrLeuTrpAsnGluMet 616
QY 1498 ATAAACATTAATCTAGCGCTACAGCGAGTCAACATAGGAGAGAGTGGTGCAGG 1557
Db 617 ValLysIleAsnProSerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArg 636
QY 1558 ATGTTGGGGATCGCTGCTGTATCGAGTGCAGCTGCTATAGATGCGGAATCCGTCAC 1617
Db 637 ValLeuGlyAspValIleAlaIleThrHisCysAlaLysIleGluGly---AsnValTyr 655
QY 1618 TTCCAAAATCTATCGAGTTATCACATCCACTCAATACATGTTATAGCCACCATTCGTT 1677
Db 656 LeuGlnAsnSerMetArgSerMetAspSer---AsnThrCysTyrSerArgProProVal 674
QY 1678 CTATTTTCATAT-----GGAGAAAACAAGGAAACATACAGGACACATCCGCTGAA 1728
Db 675 ThrPheThrIleThrLysAsnAlaAsnAsnArgGlySerIleGluGlnLeuGlyGlu 694
QY 1729 ACAACAGAGTTGCTTCCACAGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCTGTA 1788
Db 695 GluAsnGluIlePheThrGluArgLysLeuIleGluProCysAlaLeuAsnGlnLysArg 714
QY 1789 TATTTTCTGTTTGCATCCGCTTATGCTTTTAAACAACTATAATTTTGTAAATGCTA 1848
Db 715 TyrPheLysPheGlyLysGluTyrValTyrTyrGluAsnTyrThrPheValArgLysVal 734
QY 1849 GAGCGTCCCATATACAGATTGTAGCACATTTGTGAGCTTAATCTAACCCCTGCTAGAA 1908
Db 735 ProProThrGluIleGluValIleSerThrTyrValGluLeuAsnLeuThrLeuLeuGlu 754
QY 1909 GATCGGGAATTTTGGCTTTATCCGTTTACACAAAAGAGAGTGGTGTGATGTGGTGA 1968
Db 755 AspArgGluPheLeuProLeuGluValTyrThrArgAlaGluLeuGluAspThrGlyLeu 774
QY 1969 TTGATATATGACAGATAGTCCGCCAATCACTACATCACTAAATTTTATGACATA 2028
Db 775 LeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuArgPheTyrAspIle 794
QY 2029 AACAAGTATATAGAGTGGATACAAAT-----TAC 2058
Db 795 AspSerValValAsnValAsnAsnThrAlaValIleMetGlnGlyIleAlaSerPhePhe 814
QY 2059 GCGGGCTCGCAGGAATTCGCG----- 2079
Db 815 LysGlyLeuGlyLysValGlyGluAlaValGlyThrLeuValLeuGlyAlaAlaGlyAla 834
QY 2080 -----TGTATGCTCTATTACTAAAAAGATGCAAAACCAATATATGCGCAACC 2127
Db 835 ValValSerThrValSerGlyIleAlaSerPheLeuAsnAsnProPheGlyGlyLeuAla 854
QY 2127 ----- 2127
Db 855 IleGlyLeuLeuValIleAlaGlyLeuValAlaAlaPhePheAlaTyrArgTyrValMet 874
QY 2128 CAATTAGACACCGGGAATGAGTAAACAGATCTATCAATGCTAAAGCATGACATTA 2187
Db 875 GlnIleArgSerAsnProMetLysAlaLeuTyrProIleThrThrLysAlaLeuLysAsn 894
QY 2188 GCTTCACTACAAGACTATGCCAAGATTGAAGTGTATTCATCTGCTTATAGTGAAGCT 2247
Db 895 LysAlaLysThrSerTyrGlyGlnAsnGluGluAspGlySerAspPheAspGluAla 914
QY 2248 GAACAGTTTAAC-----AATAACCTTAATCAACATTAAGAA 2283
Db 915 LysLeuGluGluAlaArgGluMetIleLysTyrMetSerMetValSerAlaLeuGluLys 934
QY 2284 CAACTAAATAGCT---AAAATAATTTAGATCAGCCATCAACCAAGCTAATACGGAT 2340
Db 935 GlnGluLysLysAlaIleLysLysAsnSerGlyValGlyLeuIleAlaSerAsnValSer 954
QY 2341 AAAACGACITTTGATATGACACCCCAAAATTTAGTTGAAGCATACAAAGCACTAAAACC 2400
Db 2400
```


Db 955 LysLeuAlaLeuArgArgGlyProLysTyr-----Thr 966

Qy 2401 ACTTGAACAACGCTGCTACT 2421
 |||:|||||
 Db 967 ArgLeuGlnGlnAsnAspThr 973

RESULT 11

VGBBB2

glycoprotein B precursor - human herpesvirus 2 (strain 333)
 C:Species: human herpesvirus 2
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
 C:Accession: A26790; A45928
 R:Stuve, L.L.; Brown-Shimer, S.; Pachel, C.; Najarian, R.; Dina, D.; Burke, R.L.
 J. Virol. 61, 326-335, 1987
 A:Title: Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene
 A:Reference number: A26790; MUID:87112925; PMID:3027364
 A:Accession: A26790
 A:Molecule type: DNA
 A:Residues: 1-904 <STU>
 A:Cross-references: GB:M15118; NID:g330256; PIDN:AAA45837.1; PID:g330257
 R:Zwaagstra, J.C.; Leung, W.C.
 Can. J. Microbiol. 33, 879-887, 1987
 A:Title: The nucleotide sequence of herpes simplex virus type 2 (333) glycoprotein gB
 A:Reference number: A45928; MUID:88079667; PMID:2446730
 A:Accession: A45928
 A:Molecule type: DNA
 A:Residues: 1-34, 'AMPV', 42-307, 'T', 309-481, 'R', 483-609, 'M', 611-664, 'R', 666-904 <ZWA>
 A:Cross-references: GB:M24771; NID:g341245; PIDN:AAA60540.1; PID:g623400
 A:Note: the authors translated the codon ATG for residue 610 as Ile
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-904/Product: glycoprotein B #status predicted <MAT>
 F:724-744/Domain: transmembrane #status predicted <TM1>
 F:749-769/Domain: transmembrane #status predicted <TM2>
 F:772-792/Domain: transmembrane #status predicted <TM3>
 F:82,136,250,393,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	2.19e-93	Length:	904
Score:	1735.00	Matches:	362
Percent Similarity:	59.83%	Conservative:	143
Best Local Similarity:	42.89%	Mismatches:	269
Query Match:	30.63%	Indels:	70
DB:	1	Gaps:	16

US-09-147-052-3 (1-3261) x VGBBB2 (1-904)

Qy 58 AACATCATCTCCGAGTACCCAAAATGTGACATCAAGAGAACTTTTTCGACGCTCCAGTTC 117
 |||:|||||
 Db 82 AsnAlaThrValAlaAlaGlyHisAlaThrLeuArgAlaHisLeuArgGluIleLysVal 101

Qy 118 TCTGAGGAGAGTCTAGCTTTTATCTTTGTGCCCAACAGTGGGTCAACCGTGATCCGT 177
 |||:|||||
 Db 102 GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValValGln 121

Qy 178 CTAGAACCGCGCGAATCTCCGACCTAGAAACCCACCGAGTGGGTGAAGGAATC 237
 |||:|||||
 Db 122 PheGlnProArgCysProThrArgProGluGlnAsnTyrThrGluGlyIle 141

Qy 238 GCGATATTATTAAGAGAAATATCATGCTCCATATAAATTTAAAGTACGCTTTATATATAA 297
 |||:|||||
 Db 142 AlaValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 161

Qy 298 AATATCATTCAGAGCAGCATCGCGGGAGCAGCATATAGACATCACTAATCGATAT 357
 |||:|||||
 Db 162 AspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 181

Qy 358 ACAGATAGGACCGCTTCCATTGAAGAGATCAGCATCAATCGACGGCAAGAGA 417
 |||:|||||
 Db 182 GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaLysGlyVal 201

Qy 418 TGCTCATCTAAAGCAAGATACCTTTAGAAAACAATGTATATGTTGAAGCGCTTGACAGCGAT 477
 |||:|||||
 Db 202 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrAlaPheHisArgAsp 221

Qy 478 GCGGAGAAAAACAAGTACTTCTAAAAACCATCAAAATTCACACGCGCCGATCTAGGGCA 537
 |||:|||||
 Db 222 AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArgGly 241

Qy 538 TGGCACACGACTAATAGACAGTATACCGTGTGGGGATCACCATGGATATATCAACGGGA 597
 |||:|||||
 Db 242 TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly 261

Qy 598 ACCTCCCTCAATTTAGTAGAGGAATGATGCCGCTCTGTGTTCCGTATTCATAT 657
 |||:|||||
 Db 262 ThrThrValAsnCysIleValGluValAspAlaArgSerValTyrProTyrAspGlu 281

Qy 658 TTTCATATGGCCCAATGGCAGATCGCGAACATATCTCCATTTTATGCTATCCCAACCA 717
 |||:|||||
 Db 282 PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg 300

Qy 718 GAGCTGCCCGACAGACCCATGGGATATCCCGAGGATAATTTCAAAACAATAGATAGTAT 777
 |||:|||||
 Db 301 GlySerHisThrGluHisThrSerTyrAlaAlaAspArgPheLysGlnValAspGlyPhe 320

Qy 778 TTTTCAATGGATTTGGCAGAGCGTCCAAAGCAAGCCTTCCAGTCAAGCGTAACCTTCTC 837
 |||:|||||
 Db 321 TyrAlaArgAspLeuThrThrLysAlaAlaThrSerProThrThrArgAsnLeuLeu 340

Qy 838 ATCACAATCACATTCACAGTGGGTGGGCTGGCTCCAAAACTACTCGTGTATGTCTCA 897
 |||:|||||
 Db 341 ThrThrProLysPheThrValAlaThrPaspTyrPvalProLysArgProAlaValCysThr 360

Qy 898 ATGACTAAGTGGAGAGAGTGACTGAATGTTGGCGCAACAGTGAATGGGAGATACAGA 957
 |||:|||||
 Db 361 MetThrLysTyrGlnGluValAspGluMetLeuArgAlaGluTyrGlyCysSerPheArg 380

Qy 958 TTTATGGCCGCGAATCTTCGCAACGTTTATCAGTAATACGACGTGATGATCCCAAT 1017
 |||:|||||
 Db 381 PheSerSerAspAlaIleSerThrThrPheThrThrAsnLeuThrGlnTyrSerLeuSer 400

Qy 1018 CGCATCATATTAGGACAAATGATTAAACGCGAGGCAAGCAGCAATCGACAGATATTT 1077
 |||:|||||
 Db 401 ArgValAspLeuGlyAspCysIleGlyArgAspAlaArgGluAlaIleAspArgMetPhe 420

Qy 1078 AGGACAAAAATAATAGACAGTCCAGCTCAGCTGGTGGCATGTACATATTTCTTGCTCTC 1137
 |||:|||||
 Db 421 AlaArgLysTyrAsnAlaThrHisIleLysValGlyGlnProGlnTyrThrLeuAlaThr 440

Qy 1138 GGGGATTTATGTAGCATATCAGCTGTCTTATCCAAATCCCTGGCTCATGTACCTC 1197
 |||:|||||
 Db 441 GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnThrLeuAlaGluLeuTyrVal 460

Qy 1198 AGAGATTTGATCAGACACACAGGCCGATGAGATGCTCGACCTGGTAAACAATAAGCAT 1257
 |||:|||||
 Db 461 ArgGluTyrMetArgGluGln-----AspArgLys--- 470

Qy 1258 GCAATTTATAAGAAAAATGTACCTCATTTGTCAGAGTTGCGCGAGATATTCGAATGCA 1317
 |||:|||||
 Db 471 -----ProArgAsnAlaThrPro-----AlaProLeuArgGluAla 482

Qy 1318 CCAATAGAAAAATAACATTAGACACACACAGCATATTAATACCATGCTCTGTTCAA 1377
 |||:|||||
 Db 483 ProSerAlaAsnAlaSerValGlu-----ArgIleLysThrThrSerSerIleGlu 499

Qy 1378 TTCGCATGCTCCCAATTTCTTTATGATCATATACAAACCCATTAATATGATATGATTTAGT 1437
 |||:|||||
 Db 500 PheAlaArgLeuGlnPheThrTyrAsnHisIleGlnArgHisValAsnAspMetLeuGly 519

Qy 1438 AGGATTCACACAGCTTGGTGGAAATTCGAGAATAGAGAAGTGTGTTTATGCGCAGGAGG 1497
 |||:|||||
 Db 520 ArgIleAlaValAlaTyrCysGluLeuGlnAsnHisGluLeuThrLeuThrPasnGluAla 539

Qy 1498 ATAAGATTAATCTACGGCTACAGCGAGTGCACATCTAGGAAGAGAGTGGCTGCAAG 1557

Db	540	ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgValSerAlaArg	559
Qy	1558	ATGTTGGGGATGTCGTCTGTATCGAGCTGCAGCTCTAGATCGGAATCCGTCAC	1617
Db	560	MetLeuGlyAspValMetAlaValSerThrCysValProValAlaProAspAsnValIle	579
Qy	1618	TTGC AAAATCTATGCGAGTTATCACATCCACCTAAATACATGTTATAGCGACCATGGTT	1677
Db	580	ValGlnAsnSerMetArgValSerSerArgProGlyThrCysTyrSerArgProLeuVal	599
Qy	1678	CTATTTTCATATGGAGAAACCAAGGAAAC--ATACAGGAGCACTCGGTGAAACAAC	1734
Db	600	SerPheArgTyr---GluAspGlnGlyProLeuIleGlnGlyGlnLeuGlyGluAsnAsn	618
Qy	1735	GAGTTGCTTCCAAAGCTAGAGGCTGTAGAGCCATGCTGGCTAATCATCTAGATATTTT	1794
Db	619	GluLeuArgLeuThrArgAspAlaLeuGluProCysThrValGlyHisArgArgTyrPhe	638
Qy	1795	CTGTTTGGATCCGGTTATGCTTTATTTGAAACTATAAATTTGTTAAGTGGTAGAGCT	1854
Db	639	IlePheGlyGlyTyrValTyrPheGluGluTyrAlaTyrSerHisGlnLeuSerArg	658
Qy	1855	GCGATATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAACCTGCTAGAGATCGG	1914
Db	659	AlaAspValThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAspHis	678
Qy	1915	GAATTTTGCTTTTATCCGTTTACAAAAAGAGAGTTGCGTGATGTTGCTGATTTGGAT	1974
Db	679	GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp	698
Qy	1975	TATGCAGAAGTAGCTCCCGCAATCACTACATGAACCTAAATTTTATGACATAACAAA	2034
Db	699	TyrThrGluValGlnArgArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr	718
Qy	2035	GTAATACAAGTGGATACAAAT-----TACGCGGGGCTGCAGGNATTC-----	2076
Db	719	ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPhePheGluGly	738
Qy	2077	-----GGCTGTATGCTATTACTAAAAAAGATGCA	2106
Db	739	MetGlyAspLeuGlyArgAlaValAlaGlyLysValValMetGlyValValGlyValVal	758
Qy	2107	AACCAAAATATGGCCAAACCCAAATTAGAGCAGCGCAATGAGTTAACAGATCTAATC	2166
Db	759	SerAlaValSerGlyValSerSerPhe-----MetSerAsnProPhe	772
Qy	2167	ATGCTAAACGGATGACATTTAGCTTCACTACAGACTATCCCAAGATTGAAGCTAGTTA	2226
Db	773	GlyAlaLeuAlaValGlyLeuLeuValLeu-----AlaGlyLeu	785
Qy	2227	TCATCTCTTATAGTGAAGCTGAACAGATT-----AACATAACCTTAAATGCA	2274
Db	786	ValAlaAlaPheAlaPheArgTyrValLeuGlnLeuGlnArgAsnProMetLysAla	805
Qy	2275	-----ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAAATCAGCCATC	2322
Db	806	LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyGlyValGlyGlyGlu	825
Qy	2323	AACCAAGCTAATACGGATAAAGACGCTTTTGATATGAACACCCCAATTTAGTTGAAGCA	2382
Db	826	GlyGluGluGlyAlaGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla	843
Qy	2383	-----TACAAAGCACTAAACACCATTTCAGAA-----CAACGTGCTACT	2421
Db	844	ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluThrGluHisLysAla	863
Qy	2422	AACCTTGAAGCTTGTGCATCACTGCTTATAATCAAAATTCGCAATATTTAGTTGGATCTA	2481
Db	864	ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys	883
Qy	2482	TACAATAAAGCT	2493

Db 884 ArgAsnLysAla 887

RESULT 12

VGBEBC

glycoprotein gI precursor - bovine herpesvirus 1 (strain Colorado-1[Cooper-1])
N:Alternate names: glycoprotein 11a; glycoprotein 16; glycoprotein g130; glycoprotein
C:Contains: glycoprotein gIb; glycoprotein gIc
C:Species: bovine herpesvirus 1
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A28877; A43380
R:Whitbeck, J.C.; Bello, L.J.; Lawrence, W.C.
J. Virol. 62, 3319-3327, 1988
A:Title: Comparison of the bovine herpesvirus 1 gI gene and the herpes simplex virus
A:Reference number: A28877; MUID:88300884; PMID:2841484
A:Accession: A28877
A:Molecule type: DNA
A:Residues: 1-932 <WHI>
A:Cross-references: GB:M21474; NID:g330756; PIDN:AAA46055.1; PID:g330757
R:van Drunen Littel-van den Hurk, S.; Parker, M.D.; Fitzpatrick, D.R.; van den Hurk
Virolgie 190, 378-392, 1992
A:Title: Structural, functional, and immunological characterization of bovine herpesv
A:Reference number: A43380; MUID:92410615; PMID:1336809
A:Accession: A43380
A:Molecule type: protein
A:Residues: 68-87:505-516 <VAN>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-67/Domain: signal sequence #status predicted <SIG>
F:68-932/Product: glycoprotein gI #status predicted <GGI>
F:68-504/Product: glycoprotein gIb #status predicted <GIB>
F:505-932/Domain: glycoprotein gIc #status predicted <GIC>
F:759-827/Domain: transmembrane #status predicted <TMN>
F:105,153,441,483,640,706/Binding site: carbohydrate (Asn) (covalent) #status predict

Alignment Scores:

Pred. No.:	3,51e-93	Length:	932
Score:	1731.50	Matches:	341
Percent Similarity:	65.11%	Conservative:	105
Best Local Similarity:	49.78%	Mismatches:	218
Query Match:	30.56%	Indels:	21
DB:	1	Gaps:	7

US-09-147-052-3 (1-3261) x VGBEBC (1-932)

Qy	52	GGTACGAACTCATCTCCGAGTACCCAAATGTGACATCAAGAAAGTTGTTTCGAGCGTC	111
Db	97	GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla	116
Qy	112	CAGTTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCCCCCACCAGTGGGTTCACACCGTG	171
Db	117	GlnAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal	136
Qy	172	ATCCGTCTAGNACGCCGCCGAAATGTCGCCGACCTAGAAAGCCAGTCGGGTGAA	231
Db	137	ValArgLeuAlaProAlaAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu	156
Qy	232	GGAATCCGGATATATTATTAAGAGAAATCAGTGCCATATAATTTAAAGTGACGCTTTAT	291
Db	157	GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPhenylsAlaTyrIle---	175
Qy	292	TATAAATATCATTCAGACGACGACATGACGGGGACGACATATACAGATCATCATTAAT	351
Db	176	TyrLysAsnValIleValThrThrThrTipAlaGlySerThrTyrAlaAlaIleThrAsn	195
Qy	352	CGATATACAGTAGGACCGCGTTTCATTCAATGAAGAGATCAGGATCTAATTCGACGGA	411
Db	196	GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLysLys	215
Qy	412	GGAAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATCTATATGTTGAAGCGTTGAC	471
Db	216	TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValValAlaPheAsn	235
Qy	472	AGGATCGGGAGAAAAACAAGTACTTCTAAACCATCAAAATTTCAACGCCGCCGATCT	531

Db 98 AsnLysThrAlaGluAspAlaArgAlaGlnLeuArgGlnSerValArgGlnIleArgAla 117
QY 118 TGTGAGGAAGAGTCTACGTTTATCTTTCTCCCCACAGTGGTTCACACCGTACCGT 177
Db 118 GlnAsnAlaThrSerMetPheTyrValCysProProThrGlyAlaThrValValGln 137
QY 178 CTAGAACCCCGCGAAATGTCGCCAACCTAGAAAAGCCACCGAGTGGGGTGAAGAAATC 237
Db 138 PheGluGluProArgProCysProAspValAlaAlaGlyLysAsnPheThrGluGlyIle 157
QY 238 GCGATATTATTAAAGAGAAATACAGTCCATATAAATTTTAAAGTACGCTTTTATTATAA 297
Db 158 AlaValIlePheLysGluAsnIleAlaProTyrLysPheThrAlaThrMetTyrTyrLys 177
QY 298 ATATCATCTCAGACGACATCGGAGGGGAGCAGACATATAGACAGATCACTAATCGATAT 357
Db 178 GluIleThrValThrGlnThrProGlnGlySerArgTyrLeuGlnLeuThrGlyLeuTyr 197
QY 358 ACAGATAGACGCCGCTTCCATTGAAGAGATCAGCGATCAATCAGCGGCAAGGAAGA 417
Db 198 AsnAspArgAlaProValProPheGluGluIleThrAspValIleAsnAlaLysGlyLeu 217
QY 418 TGTCTCATCTAAAGACAGATACCTTAGAAAACATGTATATGTTGAAGCGTTTGACAGGAT 477
Db 218 CysArgSerAspValThrTyrValArgSerGlnArgValThrAlaTyrAspArgAsp 237
QY 478 GCGGGAGAAAACAGTACTCTTAACCATCAAAATTCACACGCCCGCAATCTAGGCA 537
Db 238 GluTrpGlyArgGluValLysLeuValProSerLysThrSerThrProAsnSerArgGly 257
QY 538 TGGCACACACTAATCAGACGTATACCGTGTGGGATCACCATTGGATATATCGAACGGGA 597
Db 258 TrpTyrThrAspArgMetTyrAlaProAsnAlaHisAlaGlyPheTyrLysAlaGly 277
QY 598 ACTCGCTCAATGTATAGTAGGAATAGGATCGCCGCTCTGTGTTCCGTATTCTATAT 657
Db 278 ThrThrValAsnCysIleValGluGluValGluAlaArgSerAlaTyrProTyrSerAsn 297
QY 658 TTTGCAATGGCAATCGGACATCGCGAACAATATCTCCATTTTATGGTCTATCCCCACCA 717
Db 298 PheValLeuAlaThrGlyAspPheValTyrValSerProPheTyrGlyLeuGly---Glu 316
QY 718 GAGGCTGCCGAGACCCATGGATATCCCGAGGATAATTCACAACTAGATACATAT 777
Db 317 AspAlaHisArgGluTyrAsnAlaTyrSerAlaAspArgPheLysGlnValAspGlyPhe 336
QY 778 TTTTCAATGGATTGGACAGGCTCGAAAAGCAAGCCTTCCAGCTAAGCGTAACTTCTC 837
Db 337 PheProArgAspLeuAspSerGlyGluThrAlaProGluProValValArgAsnLeuLeu 356
QY 838 ATCATCATCACTTCACAGTTGGTGGGACTGGGCTCCAAAACACTACTCGTGTATGTTCA 897
Db 357 ThrThrProLysPheThrIleGlyTrpAspTyrLysProLysAspProSerValCysSer 376
QY 898 ATGACTAAGTGAAGAGGTGACTGAAATGTTGCGTGAACACATTAATGGGAGATACAGA 957
Db 377 ValThrLysTrpGluGluValGluMetMetArgAlaGluTyrGlySerThrPheArg 396
QY 958 TTTATGCGCGGTGAATCTTCGGCAAGCTTTATCAGTAATACGACTGAGTTTGATCCAAAT 1017
Db 397 PheThrSerSerSerLeuSerAlaThrPheThrThrAsnValThrGlnTyrProGln 416
QY 1018 CGCATCATATTAGCAATGTATTAAACCGGAGGAGAAAGCAACATCGACGAGATATT 1077
Db 417 ArgIleGluLeuSerAspCysValAlaArgGluAlaGlnAlaAlaValAspAlaIleTyr 436
QY 1078 AGCACAATAATATACAGTACGTCAGGTGGACATGTACATATTCTTGGTCTC 1137
Db 437 AlaArgArgTyrAsnAlaSerHisValLysValGlyLeuGlnTyrTyrLeuAlaGln 456
QY 1138 GGGGATTTATGTAGCATATCAGCTGTTCTATCCAAATCCCTGGCTCATATGTACCTC 1197
Db 457 GlyGlyPheLeuValValTyrGlnProLeuIleSerAsnSerLeuAlaGluMetTyrLeu 476

QY 1198 AGAGAATTGATGAGAGACAACAGGACCGCATGAGATGCTCGACCTGGTAACAATAAGCAT 1257
Db 477 ArgGlu----- 478
QY 1258 GCAATTTATAAGAAAATGCTACCTCATTTGTACAGATTGGCGGAGATATTCCAAATGCA 1317
Db 479 -----AlaGluAlaArgAlaLeuGluProAlaProLeuProThrThrProAla 494
QY 1318 CCAATATAGAAAATAACATTAGACGACACACAGAGCTATTAAATCGACATCGTCTGTTCAA 1377
Db 495 ProGlu-----AlaAlaGlySerArgGlyThrLeuSerThrThrGlnSerValGlu 511
QY 1378 TTCGCCATCGCTCAATTTCTTATGATCATATACAAACCATATTAATGATATTTAGT 1437
Db 512 PheAlaArgLeuGlnPheThrTyrAspHisIleGlnLysHisValAsnGluMetLeuGly 531
QY 1438 AGSATGTCACAGCTTGGTGGCAATTCAGAAATGACAGATAGAACTGCTTTTATGGCAGAGG 1497
Db 532 ArgIleAlaAlaAlaTrpCysGlnLeuGlnAsnGlnGluValLeuTrpAsnGluAla 551
QY 1498 ATAAAGATTAACTAGCGCTACAGCGAGTGCACACATTAGGAAGAGAGTGGCTGCNAAG 1557
Db 552 ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgArgValGlyAlaArg 571
QY 1558 ATGTTGGGAGTGTCTGCTGTATCGAGCTGCTGCTATAGATCGGAATCCGTCAC 1617
Db 572 MetLeuGlyAspValMetAlaValSerThrCysIleAlaValAlaProHisAsnValIle 591
QY 1618 TTCGAAATTCATCGAGTGTATCATCCACTAATACATGTTATAGCCGACCATTTGGTT 1677
Db 592 MetGlnAsnSerMetArgLeuProAlaArgProLysThrCysTyrAlaArgProLeuVal 611
QY 1678 CTATTTTCATATGAGAAAACCAAGGAAACATACAGGACCACTCGTGAACACACGAG 1737
Db 612 SerPheArgTyrAlaAspGluGlyGluLeuIleGluGlnLeuGlyGluAspAsnGlu 631
QY 1738 TTGCTTCCACGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCGTATGATTTTCTG 1797
Db 632 IleArgLeuGluGlnAsnAsnLeuGluProCysThrValGlyHisLysArgTyrPheVal 651
QY 1798 TTTGGATCGGTTTATCTTATTTGAAACTATAATTTCTTAAGATGCTAGACGCTGCC 1857
Db 652 PheGlyAspGlyTyrValTyrPheGluGluTyrAlaTyrSerHisGlnValSerArgAla 671
QY 1858 GATATACAGATTGCTAGCATTGTCGAGCTTAATCTAACCTGCTAGAGATCGGAA 1917
Db 672 AspValProValValSerThrPheValAspLeuAsnLeuThrMetLeuGluAspHisGlu 691
QY 1918 ATTTTCCCTTATCCGTTTACAAAAGAGAGTTCGCTGATGCTTGGTGTATTTGATAT 1977
Db 692 PheLeuProLeuGlnValTyrThrArgGlnGluIleLysAspSerGlyLeuLeuAspTyr 711
QY 1978 GCAGAAGTACGTCGCCCAATCAACTACATGAACTTAATTTTATGACATAACAAAGTA 2037
Db 712 AlaGluValGlnArgAsnGlnMetHisAlaLeuArgPheSerAspIleAspHisIle 731
QY 2038 ATAGAAGTGGATACAAATTCACG-----GGGCTGCAGGAATTC 2076
Db 732 IleAsnAspThrThrAsnAlaAlaLeuMetAspGlyLeuPheArgPhe 747

RESULT 14

VGBESA

glycoprotein B precursor - simian herpesvirus SA8 (strain B264)

C:Species: simian herpesvirus SA8

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: JQ1332

R:Borchers, K.; Weigelt, W.; Buhk, H. J.; Ludwig, H.; Mankertz, J.

J. Gen. Virol. 72, 2299-2304, 1991

A:Title: Conserved domains of glycoprotein B (gB) of the monkey virus, simian agent 8

A:Reference number: JQ1332; MUID:91374035; PMID:1895066

A:Accession: JQ1332

A:Molecule type: DNA

A;Residues: 1-885 <BOR>
A;Cross-references: EMBL:X56935; NID:g60438; PIDN:CAA0256.1; PID:g60439
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-885/Product: glycoprotein B #status predicted <NAT>
F;712-732/Domain: transmembrane #status predicted <N1>
F;737-752/Domain: transmembrane #status predicted <N2>
F;760-780/Domain: transmembrane #status predicted <N3>
F;68,122,379,411,659/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;97-558,114-514,188-252,345-393,581-618/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 3,4e-92 Length: 885
Score: 1714.50 Matches: 329
Percent Similarity: 64.08% Conservativity: 108
Best Local Similarity: 48.24% Mismatches: 210
Query Match: 30.26% Indels: 35
DB: 1 Gaps: 5

US-09-147-052-3 (1-3261) x VGBESA (1-885)

QY	58	RACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTG	117
DB	68	AsnAlaSerValGluAlaGlyArgAlaThrLeuArgGluAspLeuArgGluLeuLysAla	87
QY	118	TCTAGGAAGAGTCTACGTTTATCTTGTCTCCACCAGTGGTTCACCGTGCATCGGT	177
DB	88	ArgaspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrValValGln	107
QY	178	CTAGAACCGCGCGAATGTCGCGAACCTAGAAAGCCAGCGGTGGTGAAGGATC	237
DB	108	PheGluGlnProArgProCysProArgAlaProaspGlyGlnAsnTyrThrGluGlyLe	127
QY	238	CGCATATATTAAAGAGATATCAGTCCATATAAATTAAGTGCAGCTTTATATAAA	297
DB	128	AlaValPheLysGluAsnAlaProTyrLysPheLysAlaThrMetTyrTyrLys	147
QY	298	AATATCATTCAGACGAGCATGACGGGACGACATATACAGATACATATCATAT	357
DB	148	AspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGlyLe	167
QY	358	ACAGATAGGACGCCGTTTCATTTGAAGAGATCACCGATCAATTCGACGCAAGGA	417
DB	168	GluaspArgAlaProValProPheGluGluValMetAspLysLeuAlaLysGlyVal	187
QY	418	TGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATCTTGAAGCGTTGACAGGAT	477
DB	188	CysArgSerThrAlaLysTyrValArgAsnMetGluSerThrAlaPheHisArgasp	207
QY	478	CGGGAGAAAACAGTACTTCTTAAACCATCAATAATTCACACGCCGCAATCTAGGCA	537
DB	208	AspHisGluSerAspMetAlaLeuLysProAlaLysAlaAlaThrArgThrSerArgGly	227
QY	538	TGSCACACGACTAATGAGACGTATACCGTGTGGGATCACCATGGATATATCGAACGGGA	597
DB	228	TrpHisThrThrAspLeuLysTyrAsnProAlaArgValGluAlaPheHisArgTyrGly	247
QY	598	ACCTCGCTCAATTTGTATAGTAGAGAAATGGATGCCGCTCTGTTTCGGTATTCATAT	657
DB	248	ThrThrValAsnCysLeuValGluGluValGluValGluArgSerValTyrProTyrAspGlu	267
QY	658	TTTGCAATGCCAATGGCGACATCGCGAACATATCTCCATTTTATGGTCTATCCCAACCA	717
DB	268	PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg---Asp	286
QY	718	GAGGCTGCCGCAACCCATGGATATCCCGAGGATAATTTCAACAACTAGATAGCTAT	777
DB	287	GlySerHisGlyGluHisThrAlaTyrAlaAlaAspArgPheArgGlnValAspGlyTyr	306
QY	778	TTTTCAATGGATTGGACAAGCGTTCGAAAGCAAGCTTCCAGTCAAGCGTAACTTCTC	837
DB	307	TyrGluArgAspLeuSerThrGlyArgArgAlaAlaProValThrArgAsnLeuLeu	326

QY	838	ATCACAATCACACTTCCAGTTGGGTGGGACTGGGCTCCAAAACACTACTCGTGTATGTCA	897
DB	327	ThrThrProLysPheThrValGlyTrpAlaProLysArgProSerValCysThr	346
QY	898	ATGACTAAGTGAAGAGAGTGAAGTAAATGTTGGTGGTCAACAGTTAATGGGAGATACAGA	957
DB	347	LeuThrLysTyrArgGluValAspGluMetLeuArgAlaGluTyrGlyProSerPheArg	366
QY	958	TTTATGGCCCGTGAAGTCTTCCGCAACGTTTATCATATATAGTACGATGAGTTTGGTCC	1017
DB	367	PheSerSerAlaAlaLeuSerThrThrPheThrAlaAsnArgThrGluTyrAlaLeuSer	386
QY	1018	CGCATCATATTAGGACAATGTTAAACGCCAGCAGCAGCAGCAATCGCAGATATTT	1077
DB	387	ArgValaspLeuAlaAspCysValGlyArgGluAlaArgGluAlaValaspArgilePhe	406
QY	1078	AGGACAAAATAATATGACATGACGTCAGTCAAGGTGGACATGTACAAATATTTCTTCGCTC	1137
DB	407	LeuArgArgTyrAsnGlyThrHisValLysValGlyGlnValGlnTyrTyrLeuAlaThr	426
QY	1138	GGGGATTTATTGTAGCATATACGCTGTCTTATCCAAATCCCTGGCTCATATGTACCTC	1197
DB	427	GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnAlaLeuValGluLeuTyrVal	446
QY	1198	AGAGAATTGATGAGACACAACAGCAGCGATCATGATGCTCCACCTGGTAAACAATAAGCAT	1257
DB	447	ArgGluLeuValArgGlu-----	452
QY	1258	GCAATTTTAAAGAAAATGCTACCTCATTTGTCACGATGGCGGAGATATTCGAAATGCA	1317
DB	453	-----GlnThrArgArgProAlaGlyGlyAsp	461
QY	1318	CCAAATAGAAAATAACA-----TTAGACACACACACAGCT-----ATTAAA	1359
DB	462	ProGlyGluAlaAlaThrProGlyProSerValaspProSerValGluAlaGlyLys	481
QY	1360	TCGACATCGCTCTCAATTCGCGCATGCTCCAAATTTCTTTATCATCATATACAAACCCAT	1419
DB	482	ThrThrSerSerValGluPheAlaArgLeuGlnPheThrTyrAspHisIleGlnArgHis	501
QY	1420	ATTAAATGATGTTTGTAGTGGATTCGCACAGCTGGTGGCAATTCGAGAATAGAGAACTT	1479
DB	502	ValAsnaspMetLeuGlyArgIleAlaThrAlaTyrCysGluLeuGlnAsnArgGluLeu	521
QY	1480	GTTTTATGCGACGAGGATTAAGATTAATTCCTAGCGCTACACGAGTGCACATTAGGA	1539
DB	522	ThrLeuTyrAsnGluAlaArgArgLeuAsnProGlyAlaIleAlaSerAlaThrValGly	541
QY	1540	AGGAGATGCGTGCAAGATGTTGGGGATGTCGCTGTGTATCGAGCTGCATGCTATA	1599
DB	542	ArgArgValSerAlaArgMetLeuGlyAspValMetAlaValSerThrCysValProVal	561
QY	1600	GATCGGAATCGCTACITTCGCAAAATCTATCGAGTTTATCATATCCACTACTAATCATGT	1659
DB	562	AlaProaspAsnValIleMetGlnAsnSerIleGlyValAlaAlaArgProGlyThrCys	581
QY	1660	TATAGCCGACCATGTTCTTATTTTCATATGCGAAGAAACCAAGGAAACATACAGGACNA	1719
DB	582	TyrSerArgProLeuValSerPheArgTyrGluAlaaspGlyProLeuValGluGlyGln	601
QY	1720	CTCGGTGAACACAGAGTTCCTCCACGCTAGAGGTGTAGAGCCATTCGCGCTAAT	1779
DB	602	LeuGlyGluAspAsnGluIleArgLeuGluArgaspAlaLeuGluProCysThrValGly	621
QY	1780	CATCTGTAGATATTTCTGTTGGATCCCGTTATGCTTTTATTGAAAACATAATATTTGTT	1839
DB	622	HisArgArgTyrThrPheGlyAlaGlyTyrValTyrPheGluGluTyrSer	641
QY	1840	AAGATGTTAGAGCTGCCGATATACAGATGCTAGACATATTTGTCGAGCTAACTAAC	1899
DB	642	HisGlnLeuGlyArgAlaaspValThrThrValSerThrPheIleAsnLeuLeuThr	661

QY 1900 CTGCTAGAGATCGGAAATTTTCCTTCATCCGTTTACACAAAGAGAGTTCCGTGAT 1959 Db
 Db 662 MetLeuGluAspHisGluPheValProLeuGluValThrArgGlnGluIleLysAsp 681 QY
 QY 1960 GTTGGTGTATGGATTATGCAGAAGTAGCTCGCGCAATCACTACATGAACCTTAAATTT 2019 Db
 Db 682 SerGlyLeuLeuAspTyrThrGluValGlnArgArgGlnLeuHisAlaLeuArgPhe 701 QY
 QY 2020 TATGACATAACAAAGTAATAGAAGTGGATCAAAAT-----TAGCGGGGCTGCAG 2070 Db
 Db 702 AlaAspIleAspThrValIleLysAlaAspAlaHisAlaAlaLeuPheAlaGlyLeuTyr 721 QY
 QY 2071 GAATTC 2076 Db
 Db 722 SerPhe 723 Db

RESULT 15
 VGBEK2
 glycoprotein B precursor - human herpesvirus 2 (strain HG52)
 C:Species: human herpesvirus 2
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
 C:Accession: A25611
 R:Bzik, D.J.; Debroy, C.; Fox, B.A.; Pederson, N.E.; Person, S.
 Virolgy 155, 322-333, 1986
 A:Title: The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with other herpesvirus glycoprotein genes
 A:Reference number: A25611; MUID:87071654; PMID:3024391
 A:Accession: A25611
 A:Molecule type: DNA
 A:Residues: 1-904 <B27>
 A:Cross-references: GB:M14923; NID:g330254; PIDN:AAA66440.1; PID:g330255
 R:Norais, N.; Tang, D.; Kaur, S.; Chamberlain, S.H.; Maslarsz, F.R.; Burke, R.L.; Marcus, J. Virol. 70, 7379-7387, 1996
 A:Title: Disulfide bonds of herpes simplex virus type 2 glycoprotein gB.
 A:Reference number: A58366; MUID:97048015; PMID:8892856
 A:Contents: annotation; tryptic peptide disulfide bond assignments
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-904/Product: glycoprotein B #status predicted <MAT>
 F:724-744/Domain: transmembrane #status predicted <TM1>
 F:749-769/Domain: transmembrane #status predicted <TM2>
 F:772-792/Domain: transmembrane #status predicted <TM3>
 F:82,136,250,393,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:111-570,128-326,202-266,359-407,593-630/Disulfide bonds: #status experimental

Alignment Scores:
 Pred. No.: 3,64e-92 Length: 904
 Score: 1714.00 Matches: 358
 Percent Similarity: 59.60% Conservative: 145
 Best Local Similarity: 42.42% Mismatches: 271
 Query Match: 30.26% Indels: 70
 DB: 1 Gaps: 16

us-09-147-052-3 (1-3261) x VGBEK2 (1-904)

QY 58 AACTCATCTCCGATCCCAAAATGTGACATCAAGAGAGTTGTTTCGAGCGTCAGTTG 117 QY
 Db 82 AsnAlaThrValAlaAlaGlyHisAlaThrValArgAlaHisLeuArgGluIleLysVal 101 Db
 QY 118 TCTGAGGAAGAGTCTACGTTTATCTTTGTCGCCACAGGTTGTTCAACCGTGATCCGT 177 QY
 Db 102 GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValValGln 121 Db
 QY 178 CTAGAACCGCGGAAATGTCGCGAACCTAGAAAGCCACCGAGTGGGTGAAGGAATC 237 QY
 Db 122 PheGluGlnProArgArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyIle 141 Db
 QY 238 CGCATATTATTAAAGAGATATCATGCCATATAAATTAAGTAGCGCTTTATATAA 297 QY
 Db 142 AlaValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyr-TyrLys 161 Db
 QY 298 AATATCATTCAGACGACATGACGGGGGACATATACAGATCACTAATCGATAT 357 QY

162 AspValThrValSerGlnValThrPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 181 QY
 358 ACAGATAGGAGCCGCTTCATTCAGAGATATCCAGGATCAATCGACGCGCAAGGAAGA 417 QY
 182 GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaLysGlyVal 201 Db
 418 TGCATCTCTAAGCAAGATACCTTAGAACAATATATGTTGAAGCTTTGACAGGAT 477 QY
 202 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgasp 221 Db
 478 GCGGAGAAAAAAGTACTTCTAAACCATCAAAATTTCAACGCCCGCAATATAGGGCA 537 QY
 222 AspHisGluThrAspMetGluLeuLysValAlaThrArgThrSerArgGly 241 Db
 538 TGGCACACGACTAATAGACATATACCGTGTGGGATCACCATGGATATATCGAACGGGA 597 QY
 242 TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly 261 Db
 598 ACCTCCGTCATATTGATAGAGGAATGATGCCGCTCTGTGTTCCGTATTCATAT 657 QY
 262 ThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspGlu 281 Db
 658 TTTGCAATGGCAATGGGACATCGCGAACATATCTCCATTTTATGGTCTATCCGCCCA 717 QY
 282 PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg---Glu 300 Db
 718 GAGCTGCCGAGAACCCATCGGATATCCCGAGGATAATTTCAACAACACTAGATAGCTAT 777 QY
 301 GlySerHisThrGluHisThrThrTyrAlaAlaAspArgPheLysGlnValAspGlyPhe 320 Db
 778 TTTTCATGGATTGGCAACCGTCGAAACAGCCTTCCAGTCACGCGTAACCTTCTC 837 QY
 321 TyrAlaArgAspLeuThrThrLysAlaArgAlaThrSerProThrThrArgAsnLeuLeu 340 Db
 838 ATCATCATCATCTTCACAGTTGGGTCGACGTCGCGTCCCAAACTACTCGTGATGTCA 897 QY
 341 ThrThrProLysPheThrValAlaThrPheThrValProLysArgProAlaValCysThr 360 Db
 898 ATGACTAAGTGAAGAGGTGACTGAAATGTGGTGCACAGTAATGGGAGATACAGA 957 QY
 361 MetThrLysTyrGlnGluValAspGluMetLeuArgAlaGluTyrGlySerPheArg 380 Db
 958 TTTATGGCCCTGAACTTTCGGCAACGTTTATCATAGTAATACAGCTGAGTTGATCCAAT 1017 QY
 381 PheSerSerAspAlaIleSerThrThrPheThrThrAsnLeuThrGluTyrSerLeuSer 400 Db
 1018 CGCATCATATTAGGCAATGTATTAAACGCGAGGAGCAAGCAGCATGACGAGATATTT 1077 QY
 401 ArgValAspLeuGlyAspCysIleGlyArgAspAlaArgGluAlaIleAspArgMetPhe 420 Db
 1078 AGGACAAAATATATGACAGTCACGTCAAGTTGGACATGTACAAATATTTCTTGGCTCTC 1137 QY
 421 AlaArgLysTyrAsnAlaThrHisIleLysValGlyGlnProGlnTyr-TyrGlnAlaThr 440 Db
 1138 GGGGATTTATGTAGCATATCAGCCTGTTCTATCCAAATCCCTGGCTCATATGTACCTC 1197 QY
 441 GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnThrLeuAlaGluLeuTyrVal 460 Db
 1198 AGAATATGATGAGAGACAACAGGACCATGAGATGCTGACCTGGTAAACAATAAAGCAT 1257 QY
 461 ArgGluTyrMetArgGluGln-----AspArgLys--- 470 Db
 1258 GCAATTTATAAGAAAAATGCTACCTCATTTGTACGATTGCGGCGAGATATTCGAAATGCA 1317 QY
 471 -----ProArgAsnAlaThrPro-----AlaProLeuArgGluAla 482 Db
 1318 CCAATAGAAAAATACATTAGACAGACACCATGCTATTAATCGACATCGCTGTTCAA 1377 QY
 483 ProSerAlaAsnAlaSerValGlu-----ArgIleLysThrThrSerSerIleGlu 499 Db
 1378 TTCCCATGCTCCCAATTTCTTTATGATCATATACAAACCATATTAATGATATGTTTACT 1437 QY
 500 PheAlaArgLeuGlnPheThrThrAsnHisIleGlnArgHisValAsnAspMetLeuGly 519 Db

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:02 ; Search time 5.4877 Seconds
(without alignments)
5924.998 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcaactattttagcgga.....ttctatatgttgcaaacatca 63

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlpl
-Q=/cgn2_1/usptc_spool/US09147052/runat_08102003_154341_29846/app_query.fasta_1.5980
-DB=SPREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_771@runat_08102003_154341_29846 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	115	100.0	805	12	Q98Y44 turkey herp

2	115	100.0	805	12	Q98Y45
3	115	100.0	865	12	Q83291
4	115	100.0	865	12	Q8JLW2
5	115	100.0	865	12	Q8JLW3
6	115	100.0	865	12	Q8JLW4
c	51	49.0	275	17	Q98016
7	51	44.3	352	5	O44580
8	51	44.3	791	10	Q8VXB2
9	51	44.3	336	16	Q9CM07
10	49	42.6	600	5	Q8T756
11	49	42.6	1031	5	O8I564
12	49	42.6	95	5	O8T9V8
13	48	41.7	340	4	Q98UQ1
14	48	41.7	619	10	Q8H8S5
15	48	41.7	722	5	Q8T232
16	48	41.7	858	5	Q8IL23
17	48	41.7	860	5	O27725
18	48	41.7	1388	5	O8IFM3
19	48	41.7	378	16	Q8F2B5
20	47.5	41.3	217	6	Q9N0T5
c	21	47	217	6	Q9N0T5
c	22	47	217	11	Q8C470
23	47	40.9	297	16	Q97KW6
24	47	40.9	606	5	O8T1A8
25	47	40.9	785	3	O8X0Y5
26	47	40.9	1045	5	O8T2X7
27	47	40.9	2958	5	O9VDY5
28	46	40.0	88	12	Q8B870
29	46	40.0	121	2	O87050
c	30	46	164	10	Q8W3K2
c	31	46	213	13	Q9W7E5
c	32	46	215	13	Q9DF86
c	33	46	258	17	Q974T3
34	46	40.0	264	16	Q8E8B4
35	46	40.0	418	16	Q9W297
c	36	46	418	16	Q98I62
37	46	40.0	740	8	O8W4H5
c	38	46	793	10	Q8VXB1
c	39	45.5	2792	5	O8I552
40	45	39.1	215	8	O9MFB6
41	45	39.1	318	16	O9PIB4
42	45	39.1	364	10	Q8LJF2
c	43	45	390	17	O97BC8
44	45	39.1	448	16	O9KD68
45	45	39.1	448	16	Q99TS0

ALIGNMENTS

RESULT 1

Q98Y44					
ID	Q98Y44	PRELIMINARY;	PRT;	805	AA.
AC	Q98Y44;				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Glycoprotein B (Fragment).				
OS	Turkey herpesvirus.				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Marek's disease-like viruses.				
OX	NCBI_TaxID=10390;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Davidson I.;				
RT	"Marek's disease virus - turkey isolate, qB.;"				
RL	Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF350326; AAK37557.2;				
DR	InterPro; IPR000234; Glycoprot_B.				
DR	Pfam; PF00606; Glycoprotein_B; 1.				
DR	ProDom; PD000693; Glycoprot_B; 1.				
FT	NON_TER 805				
SQ	SEQUENCE 805 AA; 91115 MW; FC8EB090895E15AA CRC64;				

Alignment Scores:

Pred. No.: 2,1e-10 Length: 805
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q98Y44 (1-805)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTTACGAC 60
Db 1 MethHstYrPheArgAsnCysIlePhePheLeuValIleLeuTyGlyThraSn 20

QY 61 TCA 63
Db 21 Ser 21

RESULT 2

Q98Y45 PRELIMINARY; PRT; 805 AA.
AC Q98Y45
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B (Fragment).
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson I.;
RT "Marek's disease virus - chicken isolate, gB gene."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350325; AAK37556.2;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
FT NON_TPR 805
SQ SEQUENCE 805 AA; 91111 MW; F88B4410A85E15AA CRC64;

Alignment Scores:

Pred. No.: 2,1e-10 Length: 805
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q98Y45 (1-805)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTTACGAC 60
Db 1 MethHstYrPheArgAsnCysIlePhePheLeuValIleLeuTyGlyThraSn 20

QY 61 TCA 63
Db 21 Ser 21

RESULT 3

Q83291 PRELIMINARY; PRT; 865 AA.
AC Q83291;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein 100 precursor.
GN GB.
OS Marek disease virus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=38013;
RN [1]

RP SEQUENCE OF 1-498 FROM N.A.

RC STRAIN-JM;
RA Sousloparov M.A., Bakhtina M.M., Krendelshtchikov A.V., Babkin I.V.;
RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
complex gp100, gp60, gp48 (b-antigen).";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X91985; CAA63039.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR KW Signal.
FT SIGNAL 1 21 POTENTIAL.

SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;
Alignment Scores:
Pred. No.: 2,1e-10 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q83291 (1-865)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTTACGAC 60
Db 1 MethHstYrPheArgAsnCysIlePhePheLeuValIleLeuTyGlyThraSn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 4

Q8JLW2 PRELIMINARY; PRT; 865 AA.
AC Q8JLW2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-attenuated GA;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (VMDV, VVMDV, vv+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY12969; RAM97702.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98020 MW; F4F526114EA02DE5 CRC64;

Alignment Scores:

Pred. No.: 2,1e-10 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q8JLW2 (1-865)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTTACGAC 60
Db 1 MethHstYrPheArgAsnCysIlePhePheLeuValIleLeuTyGlyThraSn 20

QY 61 TCA 63
 Db 21 Ser 21

RESULT 5
 Q8JLW3 PRELIMINARY; PRT; 865 AA.
 AC Q8JLW3;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N;
 RA Parcells M.S., Shamblin C.E., Dienglewiez R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 Pathotypes of Marek's Disease Viruses (VMDV, vv+MDV, vv+MDV): Mutations
 in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
 RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129967; AAM97700.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98030 MW; 9008B58B4B4E114D CRC64;

Alignment Scores:
 Pred. No.: 2.1e-10 Length: 865
 Score: 115.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q8JLW3 (1-865)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTATAGTATATCTATATGTTACGAC 60
 Db 1 MethisTyrPheArgAsnCysIlePheLeuIleValIleLeuTyrGlyThrAsn 20

QY 61 TCA 63
 Db 21 Ser 21

RESULT 6
 Q8JLW4 PRELIMINARY; PRT; 865 AA.
 AC Q8JLW4;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X;
 RA Parcells M.S., Shamblin C.E., Dienglewiez R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 Pathotypes of Marek's Disease Viruses (VMDV, vv+MDV, vv+MDV): Mutations
 in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
 RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129965; AAM97698.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.

DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98107 MW; 4D2628B5E4DEB28B CRC64;

Alignment Scores:
 Pred. No.: 2.1e-10 Length: 865
 Score: 115.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q8JLW4 (1-865)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTATAGTATATCTATATGTTACGAC 60
 Db 1 MethisTyrPheArgAsnCysIlePheLeuIleValIleLeuTyrGlyThrAsn 20

QY 61 TCA 63
 Db 21 Ser 21

RESULT 7
 Q8JLW6 PRELIMINARY; PRT; 275 AA.
 AC Q8JLW6;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein SS00308.
 GN SS00308.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; AE006665; AAK40645.1; -;
 DR InterPro; IPR006204; GHMP_Kinase.
 DR Pfam; PF00288; GHMP_Kinases; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 275 AA; 30128 MW; 228FDIA99D48C05B CRC64;

Alignment Scores:
 Pred. No.: 10 Length: 275
 Score: 51.00 Matches: 9
 Percent Similarity: 86.67% Conservative: 4
 Best Local Similarity: 60.00% Mismatches: 2
 Query Match: 49.04% Indels: 0
 DB: 17 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x Q8JLW6 (1-275)

QY 60 GTTCGTACCATAGTAATAGGAGGAAAAATATGCAATTCCG 16
 Db 52 IleLysThrIleLeuAsnTyrPheLysGluLysTyrSerIlePro 66

RESULT 8
 Q44580 PRELIMINARY; PRT; 352 AA.
 ID Q44580
 AC Q44580;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

```

Alignment Scores:
Pred. No.:          9.95          Length:          791
Score:             51.00         Matches:          10
Percent Similarity: 77.78%       Conservatives:     4
Best Local Similarity: 55.56%    Mismatches:       2
Query Match:       44.35%       Indels:           2
DB:               10           Gaps:            1

US-09-147-052-1_COPY_1_63 (1-63) x Q8VXB2 (1-791)

QY      4  CACTATTATTTAGCGGAATTCATATTTTCCCTATAGTATTCTATATGCGTAGC 57
Db      168 HisAlatATrYrYsArgAsnCysLeu-----LeuIIeValValLeuIIeGlyThr 183
          |||  :::::|||||:::  |||||:::||||  |||||

RESULT 10
Q9CM07  Q9CM07  PRELIMINARY;      PRT;      336 AA.
AC      Q9CM07
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      MGIC.
GN      MGIC OR PM1040.
OS      Pasteurella multocida.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC      Pasteurellaceae; Pasteurella.
OX      NCBI_TaxID=747;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Fm70;
RA      MEDLINE=11245866; PubMed=11248100;
RA      May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL      "Complete genomic sequence of Pasteurella multocida Fm70." ;
RL      Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR      ENBL; AEO06144; AAK03124.1; -
DR      InterPro; JPR001851; Bac_innem_transp.
DR      Pfam; PF02653; BPD_transp_2; 1.
KW      Complete Proteome.
SQ      SEQUENCE 336 AA. 35886 MW. 033DD5C766A1C152 CRC64;

```

```

Alignment Scores:
Pred. No.:      21.5          Length:      336
Score:         49.00         Matches:       7
Percent Similarity: 83.33%   Conservative: 8
Best Local Similarity: 38.89% Mismatches:    3
Query Match:    42.61%     Indels:        0
DB:             16         Gaps:           0

US-09-147-052-1_COPY1_1_63 (1-63) x Q9CM07 (1-336)

QY      1 ATGCACATATTTAGCGGAATTGCAATTTTCCCTATATAGCTATTCATATGTGGT 34
      ::: :|||:::||||| |||:::|||||:::||||| |||
Db      10 LeuaspPhePhelysglnasnAlalleIyRpheValleuleulleLeugly 27

RESULT 11
QB756 PRELIMINARY; PRT; 600 AA.
ID AC Q87756
DC OC Q87756;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBurel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBurel. 21, Last annotation update)
DE Hypothetical 68.6 kDa protein.
OS Brachnlostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Brachnlostoma.
NCBI_TaxID=7739;
RN [1]
RS SEQUENCE FROM N.A.
RA RP PubMed=11967531;
RX Abi-Rached L., Gilles A., Shina T., Pontarotti P., Inoko H.;
RL "Evidence of en bloc duplication in vertebrate genomes.";
RT Nat. Genet. 0:0-0(2002).
DR EMBL; AF391294; AAA18888.1; -.

```


OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsirlin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feidblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OJ1523_A02 genomic sequence.";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC090874; AAN08230.1; -;
 SQ SEQUENCE 619 AA: 67606 MW: E41F8B9DB774C165 CRC64;

Alignment Scores:
 Pred. No.: 31.5 Length: 619
 Score: 48.00 Matches: 8
 Percent Similarity: 71.43% Conservative: 2
 Best Local Similarity: 57.14% Mismatches: 4
 Query Match: 41.74% Indels: 0
 DB: 10 Gaps: 0
 US-09-147-052-1_COPY_1_63 (1-63) x 08H8S5 (1-619)
 Qy 21 TTGCATATTTTCCTTATGTTATCTATATGTCAGAACTC 62
 Db ::::: ||||| ||| ||| |||||
 543 ValHisLeuLeuProTyrSerIleSerSerTrpAlaGluLeu 556

Search completed: October 8, 2003, 18:11:00
 Job time : 9.4877 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:33:17 ; Search time 1.01537 Seconds
(without alignments)
5835.661 Million cell updates/sec.

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcactatttttagcgga.....ttctatatgtgtacgaactca 63

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09147052/runat_08102003_154340_29835/app_query.fasta_1.5980
-DB=SwissProt_41 -OPMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40 cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_140@runat_08102003_154340_29835 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	115	100.0	865	1 VGLB_HSVMD	P18538 marek's dis
2	48	41.7	358	1 EI24_MOUSE	Q61070 mus musculus
3	48	41.7	359	1 EI24_HUMAN	Q14681 homo sapien
C 4	47	45.2	217	1 IF4E_BOVIN	Q9n0t5 bos taurus
C 5	47	45.2	217	1 IF4E_HUMAN	P06730 homo sapien
C 6	47	45.2	217	1 IF4E_MOUSE	P20415 mus musculus
C 7	47	45.2	217	1 IF4E_RABIT	P29338 oryctolagus
8	45	39.1	135	1 DSBH_CHLMT	Q9pk17 chlamydia m
9	45	39.1	135	1 DSBH_CHLTR	O84179 chlamydia t
10	45	39.1	451	1 YQBV_BACSU	P54462 bacillus su
C 11	45	43.3	462	1 STFL_MOUSE	P33242 m steroidog
C 12	45	43.3	462	1 STFL_RAT	P50569 rattus norv
13	45	39.1	488	1 SYP_BORBU	O51363 borrelia bu
14	45	39.1	968	1 PKD2_HUMAN	Q13563 homo sapien
15	44	38.3	346	1 RDS_XENLA	Q42583 xenopus lae
16	44	38.3	361	1 MRAY_FUSNN	Q8rdq0 fusobacteri
17	44	38.3	608	1 XYNL_FIBSU	P35811 fibrobacter
18	44	38.3	655	1 NEC3_MOUSE	P29121 mus musculus

19	44	38.3	669	1 FREL_CANAL	P78588 candida alb
20	43	37.4	281	1 CCAS_CHICK	O42398 gallus gall
21	43	37.4	511	1 CP4B_HUMAN	P13584 homo sapien
22	43	37.4	1741	1 RPC1_GIALA	P25202 giardia lam
23	42	36.5	305	1 COXX_RICCN	Q921f1 rickettsia
24	42	36.5	432	1 YG3L_YEAST	P48236 saccharomyc
25	42	36.5	542	1 LNT_CHLTR	O84539 chlamydia t
26	42	36.5	590	1 NUSM_TRYBB	P04540 trypanosoma
C 27	42	40.4	600	1 NIST_LACLA	Q03203 lactococcus
28	42	36.5	679	1 YHC9_YEAST	P38738 saccharomyc
29	42	36.5	692	1 NUSC_MARPO	P06264 marchantia
30	42	36.5	855	1 XAB2_HUMAN	Q9hcs7 homo sapien
31	42	36.5	855	1 XAB2_MOUSE	Q9cdcd2 mus musculu
32	42	36.5	855	1 XAB2_RAT	Q99PK0 rattus norv
33	42	36.5	1146	1 CCAS_RAT	Q02485 rattus norv
34	42	36.5	1160	1 C1DB_BACTM	Q45747 bacillus th
35	42	36.5	1169	1 C1FB_HUMAN	O66377 bacillus th
36	42	36.5	1873	1 CCAS_HUMAN	Q13698 homo sapien
37	42	36.5	1873	1 CCAS_RABIT	P07293 oryctolagus
C 38	41.5	39.9	221	1 ERD2_PLAFA	P33948 plasmodium
39	41.5	36.1	262	1 ATP6_HAEIN	P43719 haemophilus
40	41.5	36.1	606	1 RAI7_SCHPO	P50531 schizosacch
C 41	41.5	39.9	1170	1 ITA2_BOVIN	P53710 bos taurus
42	41.5	36.1	1236	1 POLS_WEEV	P13897 western equ
43	41	35.7	111	1 YG2C_YEAST	P53445 saccharomyc
44	41	35.7	136	1 DSBH_CHLPN	Q928v7 chlamydia p
45	41	35.7	190	1 Y001_BORBU	O51035 borrelia bu

ALIGNMENTS

RESULT 1

ID	VGLB_HSVMD	STANDARD	PRT	865 AA
AC	P18538			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	GB			
OS	Marek's disease herpesvirus (strain RB-1B) (MDHV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=33707;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=89293086; PubMed=2544666;			
RA	Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;			
RT	*Nucleotide sequence and characterization of the Marek's disease			
RT	virus homologue of glycoprotein B of herpes simplex virus.*;			
RL	J. Gen. Virol. 70:1789-1804(1989).			
CC	- - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D13713; BAA02866.1; -			
DR	InterPro; IPR000234; Glycoprot_B			
DR	Pfam; PF00606; Glycoprotein_B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			
KW	Signal; Glycoprotein; Transmembrane.			
FT	SIGNAL 1 21 POTENTIAL.			
FT	CHAIN 22 865 GLYCOPROTEIN B.			
FT	DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 683 700 POTENTIAL.			
FT	TRANSMEM 709 729 POTENTIAL.			
FT	TRANSMEM 732 752 POTENTIAL.			
FT	DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).			

```
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 98091 MW; B30B93C1AC65C6C3 CRC64;

Alignment Scores:
Pred. No.: 2,36e-11 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 21
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x VGLB_HSVMD (1-865)

QY 1 ATGCACATTTTATAGCGGAATTCATATTTTCCCTATATAGTTATCTATATGGTACGAC 60
Db 1 MethIstYrPheArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrasn 20

QY 61 TCA 63
Db 21 Ser 21

RESULT 2
EI24_MOUSE
ID EI24_MOUSE STANDARD; PRT; 358 AA.
AC Q61070;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Etoposide-induced protein 2.4 (p53-Induced protein 8).
GN EI24 OR FIG8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=96226348; PubMed=8649819;
RA Lehar S.M., Nacht M., Jacks T., Vatter C.A., Chittenden T., Guild B.C.;
RT "Identification and cloning of EI24, a gene induced by p53 in
etoposide-treated cells.";
RL oncogene 12:1181-1187(1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE CELLULAR RESPONSE TO DNA DAMAGE
AND/OR P53-MEDIATED APOPTOSIS.
CC -1- TISSUE SPECIFICITY: FOUND IN ALL THE EXAMINED TISSUES. HIGH
EXPRESSION WAS FOUND IN LIVER, SKELETAL MUSCLE, PANCREAS, KIDNEY
HEART AND TO A LESSER EXTENT IN BRAIN, PLACENTA AND LUNG.
CC -1- INDUCTION: BY ETOPOSIDE TREATMENT, INDUCTION REQUIRES P53.
CC ETOPOSIDE INDUCES DNA DAMAGE IN CELLS BY INHIBITING DNA
TOPOISOMERASE II, AND ULTIMATELY CAUSES APOPTOTIC CELL DEATH.
CC -1- SIMILARITY: HIGH, TO HUMAN EI24.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U41751; AAC52483.2; -.
CC MGD; MGI:108090; EI24.
CC DOMAIN 59 62 POLY-ARG.
SQ SEQUENCE 358 AA; 40863 MW; CF690BB3FD911265 CRC64;

Alignment Scores:
Pred. No.: 5.24 Length: 358
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 1 Gaps: 0
```

```
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x EI24_MOUSE (1-358)

QY 2 TGCACATTTTATAGCGGAATTCATATTTTCCCTATATAGTTATCTATATGGTA 55
Db 93 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuLeuPheTyArgVal 110

RESULT 3
EI24_HUMAN
ID EI24_HUMAN STANDARD; PRT; 359 AA.
AC O14681;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Etoposide-induced protein 2.4 (p53-Induced protein 8).
GN EI24 OR FIG8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon cancer;
RX MEDLINE=97449378; PubMed=9305847;
RA Polyak K., Xia Y., Zweier J.L., Kinzler K.W., Vogelstein B.;
RT "A model for p53-induced apoptosis.";
RL Nature 389:300-305(1997).
CC -1- INDUCTION: By p53.
CC -1- SIMILARITY: HIGH, TO MOUSE EI24.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AF010313; AAC39531.2; -.
CC Genew; HGNC:13276; EI24.
CC MIM; 605170; -.
CC GO; GO:0006918; P:induction of apoptosis by p53; TAS.
CC DOMAIN 60 63 POLY-ARG.
SQ SEQUENCE 359 AA; 40979 MW; 5626D1F725274900 CRC64;

Alignment Scores:
Pred. No.: 5.24 Length: 359
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x EI24_HUMAN (1-359)

QY 2 TGCACATTTTATAGCGGAATTCATATTTTCCCTATATAGTTATCTATATGGTA 55
Db 94 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuLeuPheTyArgVal 111

RESULT 4
IF4E_BOVIN
ID IF4E_BOVIN STANDARD; PRT; 217 AA.
AC Q9NOT5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 4E (eIF-4E) (mRNA
cap-binding protein).
GN EIF4E.
```

OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN
 RN SEQUENCE FROM N.A.
 RA Long E., Yao J., Zhao X.;
 RT "cDNA cloning and sequence analysis of bovine eukaryotic translation
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
 CC mRNA "cap" during an early step in the initiation of protein
 CC synthesis and facilitates ribosome binding by inducing the
 CC unwinding of the mRNAs secondary structures (By similarity).
 CC -1- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
 CC (which can cycle in and out of the complex). The interaction with
 CC EIF4ENIF1 mediates the import into the nucleus (By similarity).
 CC -1- PTM: Phosphorylation increase the ability of the protein to bind
 CC to mRNA caps and to form the EIF4F complex (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib>).
 CC or send an email to license@sib-sib.ch.
 CC
 CC -----
 CC EMBL: AF257235; AAF66991.1; -
 DR HSP: P07260; IAP8.
 DR InterPro: IPR001040; TIF_EIF_4E.
 DR Pfam: PF01652; IF4E; 1.
 DR PROSITE: PD003697; TIF_EIF_4E; 1.
 DR PROSITE: PS00813; IF4E; 1.
 KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation.
 FT MOD_RES 209
 SQ SEQUENCE 217 AA; 25063 MW; DZC7D809DFE4815C CRC64;

 Alignment Scores:
 Pred. No.: 7.63 Length: 217
 Score: 47.00 Matches: 8
 Percent Similarity: 83.33% Conservative: 2
 Best Local Similarity: 66.67% Mismatches: 2
 Query Match: 45.19% Indels: 0
 DB: 1 Gaps: 0
 US-09-147-052-1_COPY_1_63 (1-63) x IF4E_BOVIN (1-217)
 Qy 39 TATAAGGAAATATGCAATTCGCCCTAAATAGTG 4
 Db 183 TyrLysGluArgLeuGlyLeuProLysIleVal 194
 RESULT 5
 IF4E_HUMAN
 ID IF4E_HUMAN STANDARD; PRT; 217 AA.
 AC P06730; Q96E95;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Eukaryotic translation initiation factor 4E (EIF4E) (mRNA
 DE cap-binding protein) (EIF-4F 25 kDa subunit).
 GN EIF4E.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87147214; PubMed=3469651;
 RA Rychlik W., Domier L.L., Gardner P.R., Hellmann G.M., Rhoads R.E.;
 RT "Amino acid sequence of the mRNA cap-binding protein from human

tissues.";
 RT Proc. Natl. Acad. Sci. U.S.A. 84:945-949(1987).
 RL [2]
 RN
 RN SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RN PARTIAL SEQUENCE.
 RX MEDLINE=91131548; PubMed=1993647;
 RA Marino M.W., Feld L.J., Jaffe E.A., Pfeiffer L.M., Han Y.-M.,
 RA Donner D.B.;
 RT "Phosphorylation of the proto-oncogene product eukaryotic initiation
 RT factor 4E is a common cellular response to tumor necrosis factor.";
 RL J. Biol. Chem. 266:2685-2688(1991).
 RN [4]
 RN MUTAGENESIS OF TRP-102; GLU-103; ASP-104 AND GLU-105.
 RX MEDLINE=911912132; PubMed=1672854;
 RA Ueda H., Iyo H., Doi M., Inoue M., Ishida T., Morioka H., Tanaka T.,
 RA Nishikawa S., Uesugi S.;
 RT "Combination of Trp and Glu residues for recognition of mRNA cap
 RT structure. Analysis of m7G base recognition site of human cap binding
 RT protein (IF-4E) by site-directed mutagenesis.";
 RL FEBS Lett. 280:207-210(1991).
 RN [5]
 RN PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=87280093; PubMed=3112145;
 RA Rychlik W., Russ M.A., Rhoads R.E.;
 RT "Phosphorylation site of eukaryotic initiation factor 4E.";
 RL J. Biol. Chem. 262:10434-10437(1987).
 RN [6]
 RN PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=93280157; PubMed=8505316;
 RA Kaufman R.J., Murtha-Riel P., Pittman D.D., Davies M.V.;
 RT "Characterization of wild-type and Ser53 mutant eukaryotic initiation
 RT factor 4E overexpression in mammalian cells.";
 RL J. Biol. Chem. 268:11902-11909(1993).
 RN [7]
 RN PHOSPHORYLATION SITE SER-53.
 RX MEDLINE=96011649; PubMed=7590282;
 RA Zhang Y., Klein H.L., Schneider R.J.;
 RT "Role of Ser-53 phosphorylation in the activity of human translation
 RT initiation factor eIF-4E in mammalian and yeast cells.";
 RL Gene 163:283-288(1995).
 RN [8]
 RN PHOSPHORYLATION SITE SER-209.
 RX MEDLINE=95301551; PubMed=7782323;
 RA Joshi B., Cai A.B., Keiper B.D., Minich W.B., Mendez R., Beach C.M.,
 RA Stepinski J., Stolarski R., Darzynkiewicz E., Rhoads R.E.;
 RT "Phosphorylation of eukaryotic protein synthesis initiation factor 4E
 RT at Ser-209.";
 RL J. Biol. Chem. 270:14597-14603(1995).
 RN [9]
 RN PHOSPHORYLATION SITE SER-209.

```

RX MEDLINE-95394927; PubMed-7665584;
RA Flynn A., Proud C.G.;
RT "Serine 209, not serine 53, is the major site of phosphorylation in
RT initiation factor eIF-4E in serum-treated Chinese hamster ovary
RT cells.";
RL J. Biol. Chem. 270:21684-21688(1995).
RN [10]
RP INTERACTION WITH EIF4ENIF1.
RC TISSUE=Fetal brain, and Placenta;
RX MEDLINE-20315895; PubMed-10856257;
RA Dostie J., Ferraiuolo M., Pause A., Adam S.A., Sonenberg N.;
RT "A novel shuttling protein, 4E-T, mediates the nuclear import of the
RT mRNA 5' cap-binding protein, eIF4E.";
RL EMBO J. 19:3142-3156(2000).
CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA "cap" during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures.
CC -1- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC (which can cycle in and out of the complex). The interaction with
CC EIF4ENIF1 mediates the import into the nucleus.
CC -1- PTM: Phosphorylation increase the ability of the protein to bind
CC to mRNA caps and to form the EIF4F complex.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC -1- CAUTION: Was originally thought to be phosphorylated on Ser-53;
CC this was latter shown to be wrong.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15353; AAC13647.1; -
DR EMBL; BC012611; AAI12611.1; -
DR PIR; A26411; A26411.
DR PDB; 1IPB; 08-MAY-02.
DR PDB; 1IFC; 08-MAY-02.
DR Genew; HGNC:3287; EIF4E.
DR GK; P06730; -
DR MIM; 133440; -
DR GO; GO:0008304; Eukaryotic translation initiation factor 4 . . . ; TAS.
DR GO; GO:0003731; F:mRNA cap binding; TAS.
DR GO; GO:0003743; F:translation initiation factor activity; TAS.
DR GO; GO:0006441; P:binding to mRNA cap; TAS.
DR InterPro; IPR001040; TIF_eIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_eIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation;
KW 3D-structure.
FT MOD_RES 209 209 PHOSPHORYLATION (BY PKC).
FT CONFLICT 127 127 D -> N (IN REF. 2).
SQ SEQUENCE 217 AA; 25037 MW; B869B8DE615E699D CRC64;

Alignment Scores:
Pred. No.: 7.63 Length: 217
Score: 47.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 45.19% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x IF4E_HUMAN (1-217)
QY 39 TATAGGAAATATGCAATTCGCGCTAAATAGTG 4
Db |||||.....:|||||.....|
183 TyrLysGluArgLeuGlyIleuProProlsileVal 194

RESULT 6
IF4E_MOUSE

```

```

ID IF4E_MOUSE STANDARD; PRT; 217 AA.
AC P20415;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 4E (eIF-4E) (mRNA
DE cap-binding protein) (eIF-4F 25 kDa subunit).
DE EIF4E.
DE
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE-89308629; PubMed-2663851;
RA Altmann M., Mueller P.P., Pelletier J., Sonenberg N., Trachsel H.;
RT "A mammalian translation initiation factor can substitute for its
RT yeast homologue in vivo.";
RL J. Biol. Chem. 264:12145-12147(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE-91244820; PubMed-2037592;
RA Jaramillo M., Pelletier J., Edery I., Nielsen P.J., Sonenberg N.;
RT eIF-4E.";
RL J. Biol. Chem. 266:10446-10451(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Rat; TISSUE=Testis;
RX MEDLINE-96109128; PubMed-8558852;
RA Miyagi Y., Kerr S., Sugiyama A., Asai A., Shibuya M., Fujimoto H.,
RA Kuchino Y.;
RT "Abundant expression of translation initiation factor EIF-4E in post-
RT meiotic germ cells of the rat testis.";
RL Lab. Invest. 73:890-898(1995).
CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA "cap" during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures.
CC -1- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
CC (which can cycle in and out of the complex). The interaction with
CC EIF4ENIF1 mediates the import into the nucleus (By similarity).
CC -1- PTM: Phosphorylation increase the ability of the protein to bind
CC to mRNA caps and to form the EIF4F complex (By similarity).
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61731; AAA37545.1; -
DR EMBL; X83399; CAA58316.1; -
DR PIR; A34295; A34295.
DR PIR; I49644; I49644.
DR PDB; 1EJ4; 15-MAR-00.
DR PDB; 1EJH; 15-MAR-00.
DR PDB; 1L8B; 12-JUN-02.
DR MGD; MGI:95305; Eif4e.
DR InterPro; IPR001040; TIF_eIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_eIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation;
KW 3D-structure.
FT MOD_RES 209 209 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT CONFLICT 70 70 E -> L (IN REF. 1).

```


SQ SEQUENCE 217 AA; 25053 MW; FC61D0FB337BCD8F CRC64;

Alignment Scores:

Pred. No.: 7.63 Length: 217
Score: 47.00 Matches: 8
Percent Similarity: 83.33% Conservatives: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 45.19% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1-copy_1_63 (1-63) x IF4E_MOUSE (1-217)

QY 39 TATAAGGAAAAATATGCAATTCGCCTAAATAGTG 4
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 183 TyrLysGIuArgLeuGluGlyLeuProLysIleVal 194

RESULT 7

ID IF4E_RABIT STANDARD; PRT; 217 AA.
AC P29338;
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Eukaryotic translation initiation factor 4E (eIF-4E) (mRNA
cap-binding protein) (eIF-4F 25 kDa subunit).
GN EIF4E.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93117125; PubMed=1475206;
RA Rychlik W., Rhoads R.E.;
RT "Nucleotide sequence of rabbit eIF-4E cDNA.";
RL Nucleic Acids Res. 20:6415-6415(1992).
CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
mRNA "cap" during an early step in the initiation of protein
synthesis and facilitates ribosome binding by inducing the
unwinding of the mRNAs secondary structures.
CC -1- SUBUNIT: EIF4F is a trimer composed of EIF4E, EIF4G and EIF4A
(which can cycle in and out of the complex). The interaction with
EIF4ENIF1 mediates the import into the nucleus.
CC -1- PTM: Phosphorylation increase the ability of the protein to bind
to mRNA caps and to form the EIF4 complex (By similarity).
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61939; CAA43943.1; -;
DR PIR; S30248; B26411.
DR HSP; P07260; IAP8.
DR InterPro; IPR001040; TIF_eIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR PRODOM; PD003697; TIF_eIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation.
FT MOD_RES 209 209 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
SQ SEQUENCE 217 AA; 25049 MW; B5A6BE12F417159 CRC64;

Alignment Scores:

Pred. No.: 7.63 Length: 217
Score: 47.00 Matches: 8
Percent Similarity: 83.33% Conservatives: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 45.19% Indels: 0

DB: 1 Gaps: 0

US-09-147-052-1-copy_1_63 (1-63) x IF4E_RABIT (1-217)

QY 39 TATAAGGAAAAATATGCAATTCGCCTAAATAGTG 4
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 183 TyrLysGIuArgLeuGluGlyLeuProLysIleVal 194

RESULT 8

DSBH_CHLMU
ID DSBH_CHLMU STANDARD; PRT; 135 AA.
AC Q9PKL7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative protein-disulfide oxidoreductase.
GN TC0448.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: Required for disulfide bond formation in some proteins
(Potential).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the dsbB family. DsbB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE002313; AAF39302.1; -;
DR PIR; B81702; B81702.
DR TIGR; TC0448; -;
DR HAMAP; MF_00287; -; 1.
DR InterPro; IPR003752; DsbB.
DR Pfam; PF02600; DsbB; 1.
DR Hypothetical protein; Electron transport; Oxidoreductase; Chaperone;
KW Redox-active center; Transmembrane; Complete proteome.
FT TRANSMEM 7 26 POTENTIAL.
FT TRANSMEM 41 60 POTENTIAL.
FT TRANSMEM 67 84 POTENTIAL.
FT TRANSMEM 109 131 POTENTIAL.
FT DISULFID 36 39 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 135 AA; 15074 MW; F0629E7E1A248277 CRC64;

Alignment Scores:

Pred. No.: 16.4 Length: 135
Score: 45.00 Matches: 8
Percent Similarity: 85.71% Conservatives: 4
Best Local Similarity: 57.14% Mismatches: 2
Query Match: 39.13% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1-copy_1_63 (1-63) x DSBH_CHLMU (1-135)

QY 7 TATTTAGGCGGAATTCATATTTTCTTATAGTTATCTTA 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 40 TyrTyrGlnArgIleCysLeuPheProLeuValIleLeu 53

```

RESULT 9
DSBH_CHLTR          STANDARD;          PRT;      135 AA.
ID   DSBH_CHLTR          AC   084179;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DE   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Putative protein-disulfide oxidoreductase.
GN   CTI176.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UN-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   [1]
CC   -!- FUNCTION: Required for disulfide bond formation in some proteins
CC   (Potential).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the dsbB family. DSBH subfamily.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
EMBL; AE001291; AAC67767.1; -
DR   PIR; G71547; G71547.
DR   HAMAP; MF_00287; -; 1.
DR   InterPro; IPR003752; DsbB.
DR   Pfam; PF02600; DsbB; 1.
KW   Hypothetical protein; Electron transport; Oxidoreductase; Chaperone;
KW   Redox-active center; Transmembrane; Complete proteome.
FT   TRANSMEM 7 26
FT   TRANSMEM 41 60 POTENTIAL.
FT   TRANSMEM 67 84 POTENTIAL.
FT   TRANSMEM 109 131 POTENTIAL.
FT   DISULFID 36 39 REDOX-ACTIVE (BY SIMILARITY).
SQ   SEQUENCE 135 AA; 15005 MW; F0FFA07B9B85E3A CRC64;

Alignment Scores:
Pred. No.:      16.4      Length:      135
Score:          45.00      Matches:      8
Percent Similarity: 85.71%      Conservative: 4
Best Local Similarity: 57.14%      Mismatches: 2
Query Match:    39.13%      Indels:      0
DB:              1      Gaps:      0

US-09-147-052-1_COPY_1_63 (1-63) x DSBH_CHLTR (1-135)
QY   7 TATTTTAGCGGAATTCATATTTTCTCTATAGTATTCTTA 48
    |||:||||| |||:|||| | |||:|||||
Db   40 TyyrTyrGlnArgIleCysLeuPheProLeuValValIleLeu 53
YQEV_BACSU
ID   YQEV_BACSU          STANDARD;          PRT;      451 AA.
AC   P54462;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DE   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yqev.
GN   YQEV.

```

```

OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=971124195; PubMed=8969508;
RA   Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA   Kobayashi Y.;
RT   "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT   the Bacillus subtilis genome containing the skin element and many
RT   sporulation genes.";
RL   Microbiology 142:3103-3111(1996).
CC   [2]
CC   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=97175542; PubMed=9023197;
RA   Homuth G., Masuda S., Mogk A., Kobayashi Y., Schumann W.;
RT   "The dnaK operon of Bacillus subtilis is heptacistronic.";
RL   J. Bacteriol. 179:1153-1164(1997).
CC   [3]
CC   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Chouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Ciolek S.K., Codani J.J., Conneton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA   Gulseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivoita C., Roche B., Rose M., Sadale Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
EMBL; D84432; BAA12468.1; -
DR   EMBL; D83717; BAA12080.1; -
DR   EMBL; Z99117; CAB14485.1; -
DR   PIR; E69952; E69952.
DR   Subtilist; BGL1646; yqev.
DR   InterPro; IPR006638; Elp3.
DR   InterPro; IPR006467; MlaB_like_C.
DR   InterPro; IPR002792; TRAM.

```


polycystic kidney disease (PKD2).";
[Am. J. Hum. Genet. 61:547-555(1997).]
[7]
RN VARIANT ADPKD PRO-356, AND VARIANT PRO-28.
RP MEDLINE=99340490; PubMed=10411676;
RX Torra R., Viribay M., Telleria D., Badenas C., Watson M., Harris P.C.,
RA Darnell A., San Millan J.L.;
RN "Seven novel mutations of the PKD2 gene in families with autosomal
RT dominant polycystic kidney disease.";
RL Kidney Int. 56:28-33(1999).
[8]
RN VARIANTS ADPKD ILE-479 DEL; 504-ARG--VAL-512 DEL AND TYR-684 DEL.
RX MEDLINE=20296613; PubMed=10835625;
RN Watnick T.J., He N., Wang K., Liang Y., Parfrey P., Hefferton D.,
RA St George-Hyslop P., Germino G.G., Pei Y.;
RN "Mutations of PKD1 in ADPKD2 cysts suggest a pathogenic effect of
RT trans-heterozygous mutations.";
RL Nat. Genet. 25:143-144(2000).
CC -1- FUNCTION: Functions as a calcium permeable cation channel. PKD1
CC AND PKD2 may function through a common signaling pathway that is
CC necessary for normal tubulogenesis.
CC -1- SUBUNIT: Interacts with PKD1. PKD1 requires the presence of PKD2
CC for stable expression. Interacts with CD2AP.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Strongly expressed in ovary, fetal and adult
CC kidney, testis, and small intestine. Not detected in peripheral
CC leukocytes.
CC -1- DOMAIN: THE C-TERMINAL IS IMPLICATED IN THE INTERACTION WITH PKD1.
CC -1- DISEASE: Defects in PKD2 are the cause of autosomal dominant
CC polycystic kidney disease type II (ADPKD) [MIM:173900]; which
CC represent approximately 15% of cases of autosomal dominant
CC polycystic kidney disease, a common autosomal dominant genetic
CC disease affecting about 1 out 1000 individuals. ADPKD is
CC characterized by progressive formation and enlargement of cysts in
CC both kidneys, typically leading to end-stage renal disease in
CC adult life. Cysts also occurs in the liver and other organs. All
CC mutations, scattered between exons 1 and 11, result in a truncated
CC PKD2 that lacks both the calcium-binding EF-hand domain and the
CC two cytoplasmic domains required for the interaction of PKD2 with
CC PKD1 and with itself. ADPKD type II is clinically milder than
CC ADPKD type I, but it has a deleterious impact on overall life
CC expectancy.
CC -1- SIMILARITY: BELONGS TO THE POLYCYSTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch).

DR EMBL; U50928; AAC50520.1; -;
DR EMBL; AF004873; AAC16004.1; -;
DR EMBL; AF004859; AAC16004.1; JOINED.
DR EMBL; AF004860; AAC16004.1; JOINED.
DR EMBL; AF004861; AAC16004.1; JOINED.
DR EMBL; AF004862; AAC16004.1; JOINED.
DR EMBL; AF004863; AAC16004.1; JOINED.
DR EMBL; AF004864; AAC16004.1; JOINED.
DR EMBL; AF004865; AAC16004.1; JOINED.
DR EMBL; AF004866; AAC16004.1; JOINED.
DR EMBL; AF004867; AAC16004.1; JOINED.
DR EMBL; AF004868; AAC16004.1; JOINED.
DR EMBL; AF004869; AAC16004.1; JOINED.
DR EMBL; AF004870; AAC16004.1; JOINED.
DR EMBL; AF004871; AAC16004.1; JOINED.
DR EMBL; AF004872; AAC16004.1; JOINED.
DR EMBL; U56813; AAC50933.1; -;
DR PIR; G02640; G02640.
DR Genew; HGNC:9009; PKD2.
DR MIM; 173910; -;
DR MIM; 173900; -;

Score:	45.00	Matches:	7
Percent Similarity:	75.00%	Conservative:	2
Best Local Similarity:	58.33%	Mismatches:	3
Query Match:	39.13%	Indels:	0
DB:	1	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x SYP_BORBU (1-488)

QY 27 ATTTTCCTATAGTATTCTATATGGTACGAACTC 62
 ||| ||||| ||||| ||||| ||||| : : : : :
 Db 37 IleMetProTyGlyTySerIleTrpSerLysIle 48

RESULT 14
 PKD2_HUMAN
 ID PKD2_HUMAN STANDARD; PRT; 968 AA.
 AC Q13563; O60441; Q15764;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE Polycystin 2 (Autosomal dominant polycystic kidney disease type II
 protein) (Polycystin) (R48321).
 GN PKD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE-96243133; PubMed-8650545;
 RA Mochizuki T., Wu G., Hayashi T., Xenophontos S.L., Veldhuisen B.,
 RA Saris J.J., Reynolds D.M., Cai Y., Gabow P.A., Pierides A.,
 RA Kimberling W.J., Breuning M.H., Deltas C.C., Peters D.J.M., Somlo S.;
 RT "PKD2, a gene for polycystic kidney disease that encodes an integral
 RT membrane protein.";
 RL Science 272:1339-1342(1996).
 RL [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE-97432829; PubMed-9286709;
 RA Hayashi T., Mochizuki T., Reynolds D.M., Wu G., Cai Y., Somlo S.;
 RT "Characterization of the exon structure of the polycystic kidney
 RT disease 2 gene (PKD2).";
 RL Genomics 44:131-136(1997).
 RL [3]
 RP SEQUENCE OF 361-968 FROM N.A.
 RP TISSUE=Breast;
 RC MEDLINE-97124839; PubMed-8954772;
 RA Schneider M.C., Rodriguez A., Nomura H., Zhou J., Morton C.C.,
 RA Reeder S.T., Wernowicz S.;
 RT "A gene similar to PKD1 maps to chromosome 4q22: a candidate gene for
 RT PKD2.";
 RL Genomics 38:1-4(1996).
 RL [4]
 RP INTERACTION WITH CD2AP.
 RX MEDLINE-20490726; PubMed-10913159;
 RA Lehtonen S., Ora A., Oikarinen V.M., Geng L., Zerial M., Somlo S.,
 RA Lehtonen E.;
 RT "In vivo interaction of the adapter protein CD2-associated protein
 RT with the type 2 polycystic kidney disease protein, polycystin-2.";
 RL J. Biol. Chem. 275:32888-32893(2000).
 RL [5]
 RP REVIEW.
 RX MEDLINE-21555829; PubMed-11698076;
 RA Stayner C., Zhou J.;
 RT "Polycystin channels and kidney disease.";
 RL Trends Pharmacol. Sci. 22:543-546(2001).
 RL [6]
 RP VARIANT ADPKD GLY-414.
 RX MEDLINE-97465600; PubMed-9326320;
 RA Veldhuisen B., Saris J.J., de Haij S., Hayashi T., Reynolds D.M.,
 RA Mochizuki T., Elles R., Fossdal R., Bogdanova N., van Dijk M.A.,
 RA Coto E., Ravine D., Noerby S., Verellen-Dumoulin C., Breuning M.H.,
 RA Somlo S., Peters D.J.M.;
 RT "A spectrum of mutations in the second gene for autosomal dominant

DR GO: GO:0015629; C:actin cytoskeleton; TAS.
 DR GO: GO:0016021; C:integral to membrane; TAS.
 DR GO: GO:0005866; C:plasma membrane; TAS.
 DR GO: GO:0005509; F:calcium ion binding activity; TAS.
 DR GO: GO:0008092; F:cytoskeletal protein binding activity; TAS.
 DR GO: GO:0008022; F:protein C-terminus binding activity; TAS.
 DR GO: GO:0005247; F:voltage-gated chloride channel activity; TAS.
 DR GO: GO:0005248; F:voltage-gated sodium channel activity; TAS.
 DR GO: GO:0006812; P:cell-matrix adhesion; TAS.
 DR GO: GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO: GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro: IPR001682; Ca/Na_pore.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR005820; M_channel_nlg.
 DR InterPro: IPR00434; PKD_1.
 DR InterPro: IPR003915; PKD_2.
 DR Pfam: PF00036; ehand; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR PRINTS: PR00500; POLYCYSTIN1.
 DR PRINTS: PR01433; POLYCYSTIN2.
 DR PROSITE: PS00018; EF_HAND; FALSE_NEG.
 DR Ionic channel; Glycoprotein; Coiled coil; Transmembrane;
 DR Calcium-binding; Disease mutation; Polymorphism.
 KW DOMAIN 1 223 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 224 244 POTENTIAL.
 FT DOMAIN 245 468 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 469 489 POTENTIAL.
 FT DOMAIN 490 505 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 506 526 POTENTIAL.
 FT DOMAIN 527 550 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 551 571 POTENTIAL.
 FT DOMAIN 572 598 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 599 639 POTENTIAL.
 FT DOMAIN 620 638 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 639 679 POTENTIAL.
 FT DOMAIN 680 968 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 95 107 POLY-GLU.
 FT DOMAIN 153 157 POLY-ARG.
 FT SITE 316 328 POLYCYSTIN MOTIF.
 FT CA_BIND 763 774 EF_HAND (POTENTIAL).
 FT DOMAIN 763 796 COILED COIL (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 362 362 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 375 375 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 28 28 R -> P (frequent polymorphism).
 FT VARIANT 356 356 /FTID-VAR_011072.
 FT VARIANT 414 414 A -> P (in ADPKD).
 FT VARIANT 414 414 W -> G (in ADPKD).
 FT VARIANT 452 452 /FTID-VAR_009195.
 FT VARIANT 479 479 I -> V (in dbSNP:1801612).
 FT VARIANT 479 479 Missing (in ADPKD; somatic mutation).
 FT VARIANT 504 512 /FTID-VAR_011074.
 FT VARIANT 512 512 Missing (in ADPKD; somatic mutation).
 FT VARIANT 584 584 /FTID-VAR_011075.
 FT VARIANT 584 584 Missing (in ADPKD; somatic mutation).
 FT CONFLICT 45 45 R -> G (in REF. 1).
 FT CONFLICT 449 449 G -> V (in REF. 3).
 SQ SEQUENCE 968 AA; 109790 MW; C23B597676CE637C CRC64;

Alignment Scores:

Pred. No.:	Score:	Length:
968	17.4	968
Matches:	45.00	9
Conservative:	72.22%	4
Mismatches:	50.00%	3
Indels:	39.13%	2
Gaps:	1	1

US-09-147-052-1_COPY_1_63 (1-63) x PKD2_HUMAN (1-968)
 QY 1 ATGCACTATTATAGCGG-----AATGCATATTTTCTTATAGTATTCTTA 48
 Db 500 LeuHisTyrPheArgSerPheTrpAsnCysLeuAspValValValValLeu 517
 RESULT 15
 RDS_XENLA
 ID RDS_XENLA STANDARD; PRT; 346 AA.
 AC 042583;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Peripherin (Retinal degeneration slow protein) (XRDS38).
 GN RDS38.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97081973; PubMed=8923216;
 RA Kedziarski W., Moghrabi W.N., Allen A.C., Jablonski-Stiemke M.M.,
 RA Azarian S.M., Bok D., Travis G.H.;
 RT "Three homologs of rds/peripherin in xenopus laevis photoreceptors
 that exhibit covalent and non-covalent interactions.";
 RL J. Cell Sci. 109:2551-2560(1996).
 CC -!- FUNCTION: MAY FUNCTION AS AN ADHESION MOLECULE INVOLVED IN
 CC STABILIZATION AND COMPACTION OF OUTER SEGMENT DISKS OR IN THE
 CC MAINTENANCE OF THE CURVATURE OF THE RIM. IT IS ESSENTIAL FOR DISK
 CC MORPHOGENESIS (BY SIMILARITY).
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: FOUND IN BOTH ROD AND CONE PHOTORECEPTORS.
 CC SPECIFICALLY IN THE RIMS AND INCISURES OF ROD AND CONE OUTER
 CC SEGMENT DISCS.
 CC -!- SIMILARITY: BELONGS TO THE RDS(PERIPHERIN) / ROM1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L79915; AAB64233.1; --
 DR InterPro: IPR000830; RDS_ROM.
 DR Pfam: PF00335; transmembrane4; 1.
 DR PRINTS: PR00218; PERIPHERNRDS.
 DR PROSITE: PS00930; RDS_ROM1.1.
 KW Vision; Cell adhesion; Photoreceptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 25 43 POTENTIAL.
 FT DOMAIN 44 61 LUMENAL (POTENTIAL).
 FT TRANSMEM 62 80 POTENTIAL.
 FT DOMAIN 81 99 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 100 123 POTENTIAL.
 FT DOMAIN 124 264 LUMENAL (POTENTIAL).
 FT TRANSMEM 265 290 POTENTIAL.
 FT DOMAIN 291 346 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 229 229 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 346 AA; 39308 MW; 0A0FE7D2A540EAAE CRC64;

Alignment Scores:
 Pred. No.: 24.9 Length: 346
 Score: 44.00 Matches: 8
 Percent Similarity: 66.67% Conservative: 4
 Best Local Similarity: 44.44% Mismatches: 6
 Query Match: 38.26% Indels: 0

DB: 1 Gaps: 0
 US-09-147-052-1_COPY_1_63 (1-63) x RDS_XENLA (1-346)
 Qy 2 TGCACATATTTAGCGCGAATGCGATATTTTCCTATATAGTATCTATATGGTA 55
 ||| ::||| ||||| ||||| ::|||
 Db 26 CysCysValLeuAlaGlyIleAlaLeuPheSerMetGlyValPheLeuLysIle 43

Search completed: October 8, 2003, 17:55:26
 Job time : 3.01537 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:56:57 ; Search time 0.704304 Seconds
(without alignments)
5245.316 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcactatttagcgga.....ttctatatgtacgaactca 63

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 137520 seqs, 29319821 residues

Total number of hits satisfying chosen parameters: 275040

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlpl
-Q=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154343_29933/app_query.fasta_1.5980
-DB=Pending_Patents_AA_New -OFFMT=rapn -SUFFIX=rapn -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_12.erunat_08102003_154343_29933 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending_Patents_AA_New:*
2: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp:*
3: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp:*
4: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp:*
5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp:*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	41.7	737	US-10-425-114A-64605	Sequence 64605, A
2	47	45.2	214	US-60-485-450-1100	Sequence 1100, Ap
3	47	45.2	217	US-60-479-073-6	Sequence 6, Appl
4	45	43.3	235	US-09-897-516A-8006	Sequence 8006, Ap
5	45	39.1	281	US-09-897-516A-8246	Sequence 8246, Ap
6	45	43.3	283	PCT-US03-21266-1	Sequence 1, Appl
7	45	43.3	462	PCT-US03-21266-7	Sequence 7, Appl
8	44	38.3	398	PCT-US03-19153-210	Sequence 210, App
9	43.5	37.8	375	US-09-674-546A-2365	Sequence 2365, Ap
10	43.5	37.8	375	US-09-674-546A-2367	Sequence 2367, Ap
11	43	37.4	488	PCT-US03-28227-5007	Sequence 5007, Ap

12	43	37.4	510	7	US-60-499-964-627	Sequence 627, App
13	43	37.4	1333	1	PCT-US03-22467-12	Sequence 12, Appl
14	42.5	37.0	255	6	US-10-425-114A-66085	Sequence 66085, A
15	42.5	37.0	375	5	US-09-674-546A-2363	Sequence 2363, Ap
16	42	36.5	88	6	US-10-425-114A-44290	Sequence 44290, A
17	42	36.5	322	6	US-10-425-114A-51002	Sequence 51002, A
18	42	36.5	365	6	US-10-425-114A-56634	Sequence 56634, A
19	42	36.5	470	6	US-10-425-114A-68776	Sequence 68776, A
20	42	36.5	884	1	PCT-US03-28227-5394	Sequence 5394, Ap
21	42	36.5	889	1	PCT-US03-28227-5393	Sequence 5393, Ap
22	42	36.5	891	1	PCT-US03-28227-5392	Sequence 5392, Ap
23	42	36.5	917	1	PCT-US03-28227-5391	Sequence 5391, Ap
24	42	36.5	927	1	PCT-US03-28227-5390	Sequence 5390, Ap
25	42	36.5	1787	6	US-10-357-885-30	Sequence 30, Appl
26	42	36.5	1873	6	US-10-357-885-28	Sequence 28, Appl
27	41.5	36.1	487	1	PCT-US03-26491-250	Sequence 250, App
28	41.5	36.1	487	6	US-10-648-593-250	Sequence 250, App
C 29	41	39.4	79	6	US-10-425-114A-68269	Sequence 68269, A
30	41	39.4	82	6	US-10-425-114A-46211	Sequence 46211, A
C 31	41	39.4	97	6	US-10-425-114A-68846	Sequence 68846, A
32	41	35.7	127	6	US-10-425-114A-57106	Sequence 57106, A
33	41	35.7	144	6	US-10-425-114A-52902	Sequence 52902, A
34	41	35.7	311	6	US-10-343-650A-260	Sequence 260, App
35	41	35.7	311	6	US-10-343-650A-364	Sequence 364, App
36	41	35.7	311	7	US-60-500-315-1988	Sequence 1988, Ap
37	41	35.7	329	1	PCT-US03-18840-112	Sequence 112, App
38	41	35.7	329	1	PCT-US03-18840-182	Sequence 182, App
39	41	35.7	457	6	US-10-425-114A-55737	Sequence 55737, A
40	41	35.7	460	6	US-10-425-114A-63466	Sequence 63466, A
41	41	35.7	464	6	US-10-425-114A-55396	Sequence 55396, A
42	41	35.7	465	6	US-10-425-114A-39344	Sequence 39344, A
43	41	35.7	470	6	US-10-425-114A-73096	Sequence 73096, A
44	41	35.7	470	7	US-60-498-183-152	Sequence 152, App
C 45	41	39.4	577	6	US-10-425-114A-39598	Sequence 39598, A

ALIGNMENTS

RESULT 1
US-10-425-114A-64605
; Sequence 64605, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64605
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4569-009-H11_FLI pep
US-10-425-114A-64605

Alignment Scores:
Pred. No.: 2.88 Length: 737
Score: 48.00 Matches: 10
Percent Similarity: 77.78% Conservative: 4
Best Local Similarity: 55.56% Mismatches: 2
Query Match: 41.74% Indels: 2
DB: 6 Gaps: 1

US-09-147-052-1_COPY_1_63 (1-63) x US-10-425-114A-64605 (1-737)

QY 4 CACTATTTAGCGGAATTGCATATTTTCCCTATAGTTATCTATATGTTAGC 57

```

Db      119 HisGlnPheargLysasnValIle-----LeuIleuValleuPheGlyThr 134
      |||  |||||:|||| |||  |||||:||||:||||:|||||
RESULT 2
US-60-485-450-1100
; Sequence 1100, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1100
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-485-450-1100

Alignment Scores:
Pred. No.:      3.81      Length:      214
Score:          47.00      Matches:      8
Percent Similarity: 83.33%  Conservative: 2
Best Local Similarity: 66.67%  Mismatches: 2
Query Match:    45.19%      Indels:      0
DB:             7          Gaps:          0

US-09-147-052-1_COPY_1_63 (1-63) x US-60-485-450-1100 (1-214)
QY      39 TATAAGGAAAAATATGCAATTCGCCTAAATAAGTG 4
      |||||:||||: |||||:||||:||||:||||:||||:
Db      180 TyrLysGluargLeuGlyLeuProProlLysIleVal 191
RESULT 3
US-60-479-073-6
; Sequence 6, Application US/60479073
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Saunders, Michael John Scott
; APPLICANT: Logghe, Marc Georges
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; TITLE OF INVENTION: sequences encoding such amino acid sequences.
; FILE REFERENCE: D00590.70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-479-073-6

Alignment Scores:
Pred. No.:      3.81      Length:      217
Score:          47.00      Matches:      8
Percent Similarity: 83.33%  Conservative: 2
Best Local Similarity: 66.67%  Mismatches: 2
Query Match:    45.19%      Indels:      0
DB:             7          Gaps:          0

US-09-147-052-1_COPY_1_63 (1-63) x US-60-479-073-6 (1-217)
QY      39 TATAAGGAAAAATATGCAATTCGCCTAAATAAGTG 4
      |||||:||||: |||||:||||:||||:||||:||||:
Db      183 TyrLysGluargLeuGlyLeuProProlLysIleVal 194

```

```
QY 7 TATTTAGCGGAATTCGATA-----TTTTCCCTTATA----- 39
Db 93 TyrPheArgArgCysTyrGlnPheAlaGluPheValLeuLeuLeuCysGlyAla 112
QY 40 ---GTTATTCTATATGGTACGAAC 60
Db 113 AlaValLeuPheTyrGlyIleAsn 120

RESULT 6
PCT-US03-21266-1
; Sequence 1, Application PC/TUS0321266
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; APPLICANT: Ingraham, Holly A.
; TITLE OF INVENTION: Steroidogenic Factor-1 Protein Variants
; FILE REFERENCE: 66778-122
; CURRENT APPLICATION NUMBER: PCT/US03/21266
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US 60/395,371
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US03-21266-1

Alignment Scores:
Pred. No.: 8.62 Length: 283
Score: 45.00 Matches: 9
Percent Similarity: 68.75% Conservative: 2
Best Local Similarity: 56.25% Mismatches: 5
Query Match: 43.27% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x PCT-US03-21266-1 (1-283)
QY 57 CTTACCATATAGTAACATAAGGAAAAATATGCAATTCGCGCTAAA 10
Db 13 ArgThrIleLysSerGluTyrProGluProTyrAlaSerProProGln 28

RESULT 7
PCT-US03-21266-7
; Sequence 7, Application PC/TUS0321266
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; APPLICANT: Ingraham, Holly A.
; TITLE OF INVENTION: Steroidogenic Factor-1 Protein Variants
; FILE REFERENCE: 66778-122
; CURRENT APPLICATION NUMBER: PCT/US03/21266
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US 60/395,371
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US03-21266-7

Alignment Scores:
Pred. No.: 9.03 Length: 462
Score: 45.00 Matches: 9
Percent Similarity: 68.75% Conservative: 2
Best Local Similarity: 56.25% Mismatches: 5
Query Match: 43.27% Indels: 0
DB: 0 Gaps: 0
```

```
DB: 1 Gaps: 0
US-09-147-052-1_COPY_1_63 (1-63) x PCT-US03-21266-7 (1-462)
QY 57 CTTACCATATAGTAACATAAGGAAAAATATGCAATTCGCGCTAAA 10
Db 191 ArgThrIleLysSerGluTyrProGluProTyrAlaSerProProGln 206

RESULT 8
PCT-US03-19153-210
; Sequence 210, Application PC/TUS0319153
; GENERAL INFORMATION:
; APPLICANT: Diversa Corporation
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND
; FILE REFERENCE: 09010-290W01
; CURRENT APPLICATION NUMBER: PCT/US03/19153
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 378
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)...(25)
PCT-US03-19153-210

Alignment Scores:
Pred. No.: 13.2 Length: 398
Score: 44.00 Matches: 6
Percent Similarity: 77.78% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 38.26% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x PCT-US03-19153-210 (1-398)
QY 33 CCTTATAGTTATTCATATATGATGACGAA 59
Db 327 ProTyrHisTyrGluIleTyrTyrGln 335

RESULT 9
US-09-674-546A-2365
; Sequence 2365, Application US/09674546A
; GENERAL INFORMATION:
; APPLICANT: Institute for Genomic Research
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions
; FILE REFERENCE: CHIR-0334
; CURRENT APPLICATION NUMBER: US/09/674,546A
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 3264
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2365
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-674-546A-2365

Alignment Scores:
Pred. No.: 16 Length: 375
Score: 43.50 Matches: 10
Percent Similarity: 60.00% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 37.83% Indels: 3
DB: 5 Gaps: 1
```

755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: October 8, 2003, 17:53:53 ; Search time 19.0632 Seconds
(without alignments)
6014.205 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63
Perfect score: 115
Sequence: 1 atgcactattttaggcgaa.....ttctatatgttgtaagaactca 63

Scoring table:
BLOSOM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 11457514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_Spool/US09147052/runat_08102003_154343_29901/app_query.fasta_1.5980
-DB=Pending_Patents_AA_Main -QFMT=fastan -SUFFIX=rapm -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_2602 @runat_08102003_154343_29901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_AA_Main.*

- 1: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*

- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US106_COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	62	27	US-10-131-591A-5
2	115	100.0	456	15	US-09-147-052-2
3	115	100.0	456	24	US-09-901-572A-3
4	115	100.0	865	3	US-07-722-860-13
5	115	100.0	865	6	US-08-213-449A-13
6	115	100.0	865	6	US-08-213-449B-13
7	115	100.0	1086	15	US-09-147-052-4
8	109	94.8	62	27	US-10-131-591A-6
9	58	50.4	54	18	US-09-450-969-5904
10	58	50.4	61	18	US-09-450-969-5109
11	58	50.4	61	26	US-10-092-411A-4538
12	53	46.1	395	27	US-10-179-131-7775
13	53	46.1	582	1	PCT-US02-03987-15297
14	53	46.1	582	26	US-10-032-585-7353
15	53	46.1	582	26	US-10-072-851-15297
16	53	46.1	582	32	US-60-314-050-7353
17	52.5	45.7	41	18	US-09-450-969-6117
18	51	44.3	50	1	PCT-US01-14827-12993
19	51	44.3	51	22	US-09-758-472-5021
20	51	44.3	51	28	US-10-235-926-5021
21	49	42.6	2026	30	US-10-437-963-192277
22	48.5	42.2	473	18	US-09-438-944-585
23	48	41.7	113	23	US-09-834-366-17164
24	48	41.7	113	32	US-60-197-873-17164
25	48	41.7	255	21	US-09-733-089-15843
26	48	41.7	255	23	US-09-816-660-15843
27	48	41.7	276	1	PCT-US01-18569-3128
28	48	41.7	276	28	US-10-264-049-3128
29	48	41.7	317	10	US-08-619-362-8
30	48	41.7	317	17	US-09-373-896-3
31	48	41.7	318	10	US-08-619-362-9
32	48	41.7	340	17	US-09-373-896-1
33	48	41.7	340	32	US-60-443-566-2431
34	48	41.7	340	32	US-60-455-444-7821
35	48	41.7	340	32	US-60-485-241-7821
36	48	41.7	342	23	US-09-809-391-695
37	48	41.7	342	23	US-09-882-171-695
38	48	41.7	342	27	US-10-164-861-695
39	48	41.7	345	24	US-09-949-016-10427
40	48	41.7	345	32	US-60-443-566-2429
41	48	41.7	345	32	US-60-455-444-7819
42	48	41.7	345	32	US-60-465-241-7819
43	48	41.7	358	29	US-10-205-219-131
44	48	41.7	358	29	US-10-385-450-8
45	48	41.7	358	29	US-10-385-450-9

ALIGNMENTS

RESULT 1

US-10-131-591A-5

- ; Sequence 5, Application US/10131591A
- ; GENERAL INFORMATION:
- ; APPLICANT: Nippon Zeon Co., Ltd.,
- ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
- ; FILE REFERENCE: J209
- ; CURRENT APPLICATION NUMBER: US/10/131,591A
- ; CURRENT FILING DATE: 2002-08-15
- ; NUMBER OF SEQ ID NOS: 79

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: MDVgB signal
US-10-131-591A-5

Alignment Scores:
Pred. No.:          Length:      62
Score:             1.07e-09      21
Percent Similarity: 100.00%      21
Best Local Similarity: 100.00%
Query Match:       100.00%
Indels:            0
DB:                27
Gaps:              0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-131-591A-5 (1-62)
QY 1 ATGCACATATTTAGCGGAATTGCATATTTTCCTTATAGTTATCTATATGATGACGAAC 60
Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrglyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 2
US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Alignment Scores:
Pred. No.:          Length:      456
Score:             1.03e-09      21
Percent Similarity: 100.00%      21
Best Local Similarity: 100.00%
Query Match:       100.00%
Indels:            0
DB:                15
Gaps:              0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-147-052-2 (1-456)
QY 1 ATGCACATATTTAGCGGAATTGCATATTTTCCTTATAGTTATCTATATGATGACGAAC 60
Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrglyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 3
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

Alignment Scores:
Pred. No.:          Length:      456
Score:             1.03e-09      21
Percent Similarity: 100.00%      21
Best Local Similarity: 100.00%
Query Match:       100.00%
Indels:            0
DB:                24
Gaps:              0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-901-572A-3 (1-456)
QY 1 ATGCACATATTTAGCGGAATTGCATATTTTCCTTATAGTTATCTATATGATGACGAAC 60
Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrglyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 4
US-07-722-860-13
; Sequence 13, Application US/07722860
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohel
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 North Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/722,860
; FILING DATE: 19910628
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-101P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID

```



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-722-860-13

Alignment Scores:
Pred. No.: 1.02e-09 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-07-722-860-13 (1-865)
QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTCTATATGTTACGAC 60
Db 1 MethIstYrPheArGArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20

QY 61 TCA 63
Db 21 Ser 21

RESULT 6
US-08-213-449B-13
; Sequence 13, Application US/08213449B
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213,449B
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-108P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-213-449B-13

Alignment Scores:
Pred. No.: 1.02e-09 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-08-213-449B-13 (1-865)
QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTCTATATGTTACGAC 60
Db 1 MethIstYrPheArGArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20

QY 61 TCA 63
Db 21 Ser 21

RESULT 7
US-09-147-052-4

```

```
; Sequence 4, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4
Alignment Scores:
Pred. No.: 1.01e-09 Length: 1086
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-147-052-4 (1-1086)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTTCTATATGATGACGAC 60
Db 1 MethisTyrPheArgAsnCysIlePhePheLeuValIleLeuTyrGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21
RESULT 8
; Sequence 6, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified VgB signal
US-10-131-591A-6
Alignment Scores:
Pred. No.: 1.05e-08 Length: 62
Score: 109.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 94.78% Indels: 0
DB: 27 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-131-591A-6 (1-62)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTTCTATATGATGACGAC 60
Db 1 MethisTyrPheArgAsnCysIlePhePheLeuValIleLeuTyrGlyThrGln 20
..
```

```
QY 61 TCA 63
Db 21 Ser 21
RESULT 9
US-09-450-969-5904
; Sequence 5904, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 5904
; LENGTH: 54
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-5904
Alignment Scores:
Pred. No.: 2.94 Length: 54
Score: 58.00 Matches: 9
Percent Similarity: 82.35% Conservative: 5
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 50.43% Indels: 0
DB: 18 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-450-969-5904 (1-54)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTTCTATATGATGACGAC 51
Db 30 LeuTyrTyrPheAsnArgAsnCysLeuPheLeuLeuLeuLeuLeuLeuPhe 46
RESULT 10
US-09-450-969-5109
; Sequence 5109, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 5109
; LENGTH: 61
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-5109
Alignment Scores:
Pred. No.: 2.93 Length: 61
Score: 58.00 Matches: 9
Percent Similarity: 81.25% Conservative: 4
Best Local Similarity: 56.25% Mismatches: 3
Query Match: 50.43% Indels: 0
DB: 18 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-450-969-5109 (1-61)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTTCTATATGATGACGAC 48
Db 5 LeuHisTyrPheAsnArgAsnCysLeuTyrLeuLeuLeuLeuLeuLeuLeu 20
RESULT 11
US-10-092-411A-4538
; Sequence 4538, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4538
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4538

Alignment Scores:
Pred. No.: 2.93 Length: 61
Score: 58.00 Matches: 9
Percent Similarity: 81.25% Conservative: 2
Best Local Similarity: 56.25% Mismatches: 3
Query Match: 50.43% Indels: 0
DB: 26 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-092-411A-4538 (1-61)
Qy 1 ATGCATATTTTTCCTTATAGTATTCATATTTTTCCTTATAGTATTCAT 48
Db 5 LeuHisTyrPheAsnArgAsnCysLeuTyrLeuLeuileLeuGluLeu 20

RESULT 12
US-10-179-131-7775
; Sequence 7775, Application US/10179131
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORK
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
; FILE REFERENCE: 2976-4031
; CURRENT APPLICATION NUMBER: US/10/179,131
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 10194
; SEQ ID NO 7775
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-179-131-7775

Alignment Scores:
Pred. No.: 19 Length: 395
Score: 53.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 46.09% Indels: 0
DB: 27 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-179-131-7775 (1-395)
Qy 21 TTGCATATTTTTCCTTATAGTATTCATATTCATATTCATATTCAT 62
Db 206 LeuGlnIlePheProTyrAsnTrpSerCysTrpGlnGluLeu 219

RESULT 13
PCT-US02-03987-15297
; Sequence 15297, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elutra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits d
```

```
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15297
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Candida albicans
PCT-US02-03987-15297

Alignment Scores:
Pred. No.: 18.8 Length: 582
Score: 53.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 46.09% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x PCT-US02-03987-15297 (1-582)
Qy 21 TTGCATATTTTTCCTTATAGTATTCATATTCATATTCATATTCAT 62
Db 190 LeuGlnIlePheProTyrAsnTrpSerCysTrpGlnGluLeu 203

RESULT 14
US-10-032-585-7353
; Sequence 7353, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7353
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7353

Alignment Scores:
Pred. No.: 18.8 Length: 582
Score: 53.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 46.09% Indels: 0
DB: 26 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-032-585-7353 (1-582)
Qy 21 TTGCATATTTTTCCTTATAGTATTCATATTCATATTCATATTCAT 62
Db 190 LeuGlnIlePheProTyrAsnTrpSerCysTrpGlnGluLeu 203

RESULT 15
US-10-072-851-15297
; Sequence 15297, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
```

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15297
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-072-851-15297

Alignment Scores:
Pred. No.:      18.8      Length:      582
Score:          53.00     Matches:      9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match:     46.09% Indels:      0
DB:              26      Gaps:      0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-072-851-15297 (1-582)
QY      21 TTGCATATTTTCCTTATAGTTATCTATATGTCAGTACTC 62
      ||| ||||| ||||| ||||| ||||| ||||| |||||
Db      190 LeuGlnIlePheProTyrAsnTrpSerCysTrpGlnGluLeu 203

Search completed: October 8, 2003, 19:14:08
Job time : 21.0632 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 18:12:37 ; Search time 3.45109 Seconds
(without alignments)
5882.831 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115
Sequence: 1 atgcacatttttagcgga.....ttctatatgttgaactca 63

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 1201306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09147052/runat_08102003_154344_29960/app_query.fasta_1.5980
-DB=PublishedApplications_AA -QEXT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOPCCI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09147052 -RCGN_1_355=runat_08102003_154344_29960
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV.TIMEOUT=130 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_AA:

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

Alignment Scores:	4.05e-10	Length:	62
Pred. No.:	115.00	Matches:	21
Score:	100.00%	Conservative:	0

ALIGNMENTS

RESULT 1

US-10-131-591A-5
; Sequence 5, Application US/10131591A
; Publication No. US20030059799A1

; GENERAL INFORMATION:

; APPLICANT: Nippon Zeon Co., Ltd.,

; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof

; FILE REFERENCE: J209

; CURRENT APPLICATION NUMBER: US/10/131,591A

; CURRENT FILING DATE: 2002-08-15

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 62

; TYPE: PRT

; ORGANISM: Marek's disease gammaherpesvirus

; FEATURE:

; OTHER INFORMATION: MDVgB signal

US-10-131-591A-5

1	115	100.0	62	15	US-10-131-591A-5	Sequence 5, Appli
2	115	100.0	456	9	US-09-147-052-2	Sequence 2, Appli
3	115	100.0	456	12	US-09-901-572A-3	Sequence 3, Appli
4	115	100.0	1086	9	US-09-147-052-4	Sequence 4, Appli
5	109	94.8	62	15	US-10-131-591A-6	Sequence 6, Appli
6	53	46.1	582	12	US-10-032-585-7353	Sequence 7353, Ap
7	48	41.7	317	10	US-09-151-771-8	Sequence 8, Appli
8	48	41.7	318	10	US-09-151-771-9	Sequence 9, Appli
9	48	41.7	342	11	US-09-809-391-695	Sequence 695, App
10	48	41.7	342	12	US-09-882-171-695	Sequence 695, App
11	48	41.7	358	12	US-10-205-219-131	Sequence 131, App
12	48	41.7	358	12	US-10-385-450-8	Sequence 8, Appli
13	48	41.7	358	12	US-10-385-450-9	Sequence 9, Appli
14	47	40.9	69	14	US-10-001-876-176	Sequence 176, App
15	45	39.1	110	9	US-09-939-980-391	Sequence 391, App
16	45	39.1	968	9	US-09-753-008-7	Sequence 7, Appli
17	44.5	38.7	866	9	US-09-753-008-1	Sequence 1, Appli
18	44	38.3	202	12	US-10-307-441-18	Sequence 18, Appli
19	44	38.3	1550	10	US-09-995-542-8	Sequence 8, Appli
20	44	38.3	2100	10	US-09-995-542-6	Sequence 6, Appli
21	44	38.3	2144	9	US-09-858-194-2	Sequence 2, Appli
22	44	38.3	2144	12	US-10-154-419-2	Sequence 2, Appli
23	44	38.3	2146	10	US-09-995-542-5	Sequence 5, Appli
24	43	37.4	83	9	US-09-864-761-46530	Sequence 46530, A
25	42	36.5	97	12	US-10-319-763-183	Sequence 183, App
26	42	40.4	396	9	US-09-765-272-16	Sequence 16, Appli
27	42	40.4	419	11	US-09-769-787-157	Sequence 157, App
28	42	36.5	768	10	US-09-973-451-8	Sequence 8, Appli
29	42	40.4	845	15	US-10-242-056-37	Sequence 37, Appli
30	42	36.5	852	9	US-09-752-639-153	Sequence 153, App
31	42	36.5	852	10	US-09-984-198-153	Sequence 153, App
32	42	36.5	1338	14	US-10-029-413A-10	Sequence 10, Appli
33	42	40.4	1849	15	US-10-242-056-49	Sequence 49, Appli
34	42	36.5	1854	14	US-10-029-413A-2	Sequence 2, Appli
35	42	36.5	1873	14	US-10-029-413A-12	Sequence 12, Appli
36	42	36.5	1873	14	US-10-029-413A-14	Sequence 14, Appli
37	42	36.5	1873	14	US-10-029-413A-22	Sequence 22, Appli
38	42	40.4	2516	9	US-09-817-514A-2	Sequence 2, Appli
39	42	40.4	2516	15	US-10-242-056-47	Sequence 47, Appli
40	41.5	36.1	89	9	US-09-860-232A-9	Sequence 9, Appli
41	41.5	36.1	262	9	US-09-815-242A-9	Sequence 11030, A
42	41.5	36.1	307	12	US-10-002-631C-30	Sequence 30, Appli
43	41.5	36.1	473	11	US-09-374-046A-68	Sequence 68, Appli
44	41.5	36.1	485	15	US-10-191-398A-5	Sequence 5, Appli
45	41.5	36.1	1236	12	US-10-023-649-4	Sequence 4, Appli

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-131-591A-5 (1-62)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTTATAGTATCTATATGCTACGAAC 60
Db 1 MethisTyPheArgAsnCysIlePheLeuIleValIleLeuTyGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 2
US-09-147-052-2
; Sequence 2, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Alignment Scores:
Pred. No.: 4.51e-10 Length: 456
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-147-052-2 (1-456)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTTATAGTATCTATATGCTACGAAC 60
Db 1 MethisTyPheArgAsnCysIlePheLeuIleValIleLeuTyGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 3
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; Publication No. US2003016534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum

```

```

; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40R-S
US-09-901-572A-3

```

```

Alignment Scores:
Pred. No.: 4.51e-10 Length: 456
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

```

US-09-147-052-1_COPY_1_63 (1-63) x US-09-901-572A-3 (1-456)

```

```

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTTATAGTATCTATATGCTACGAAC 60
Db 1 MethisTyPheArgAsnCysIlePheLeuIleValIleLeuTyGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

```

RESULT 4

```

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

```

Alignment Scores:

```

Pred. No.: 4.72e-10 Length: 1086
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-147-052-1_COPY_1_63 (1-63) x US-09-147-052-4 (1-1086)

```

```

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTTATAGTATCTATATGCTACGAAC 60
Db 1 MethisTyPheArgAsnCysIlePheLeuIleValIleLeuTyGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

```

RESULT 5

```

US-10-131-591A-6
; Sequence 6, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209

```

```

; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified Vgb signal
US-10-131-591A-6

```

```

Alignment Scores:
Pred. No.: 3.79e-09 Length: 62
Score: 109.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 94.78% Indels: 0
DB: 15 Gaps: 0

```

```
US-09-147-052-1_COPY_1_63 (1-63) x US-10-131-591A-6 (1-62)
```

```

Qy 1 ATGCATATTTAGCGGAATTCATATTTTCCTTATAGTATTTCTATATGTTACGAAC 60
|||||
Db 1 MethistyrPheargArgasnCysIlePhePheLeuIleValIleLeuTyrglyThrGln 20

```

```

Qy 61 TCA 63
Db 21 Ser 21

```

RESULT 6

```
US-10-032-585-7353
; Sequence 7353, Application US/10032585
; Publication No. US20030180953A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone

```

```

; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

```

```

; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032.585

```

```

; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7353

```

```

; LENGTH: 582
; TYPE: PRT

```

```

; ORGANISM: Candida albicans
US-10-032-585-7353

```

Alignment Scores:

```

Pred. No.: 4.9 Length: 582
Score: 53.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 46.09% Indels: 0
DB: 12 Gaps: 0

```

```
US-09-147-052-1_COPY_1_63 (1-63) x US-10-032-585-7353 (1-582)
```

```

Qy 21 TTGCATATTTTCCTTATAGTATTTCTATATGTTACGAATC 62
|||||
Db 190 LeuGlnIlePheProTyrAsnTrpSerCysTrpGlnGluLeu 203

```

RESULT 7

```
US-09-151-771-8
```

```

; Sequence 8, Application US/09151771
; Publication No. US20020192745A1

```

```

; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and

```

```

; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,

```

Alignment Scores:

```

Pred. No.: 4.9 Length: 582
Score: 53.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 46.09% Indels: 0
DB: 12 Gaps: 0

```

```
US-09-147-052-1_COPY_1_63 (1-63) x US-09-151-771-8 (1-317)
```

Alignment Scores:

```

Pred. No.: 30.5 Length: 317
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 10 Gaps: 0

```

US-09-147-052-1_COPY_1_63 (1-63) x US-09-151-771-8 (1-317)

```

Qy 2 TGCATATTTAGCGGAATTCATATTTTCCTTATAGTATTTCTATATGTTACGA 55
|||||
Db 93 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuPheTyArgVal 110

```

RESULT 8

```
US-09-151-771-9
```

```

; Sequence 9, Application US/09151771
; Publication No. US20020192745A1

```

```

; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and

```

```

; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,

```

```

; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.

```

```

; CITY: Washington
; STATE: District of Columbia

```

```

; COUNTRY: U.S.A.
; ZIP: 20004

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
; NUMBER OF SEQUENCES: 17

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr

```

```

; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington

```

```

; STATE: District of Columbia
; COUNTRY: U.S.A.

```

```

; ZIP: 20004
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE

```

```

; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.

```

```

; CITY: Washington
; STATE: District of Columbia

```

```

; COUNTRY: U.S.A.
; ZIP: 20004

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/151,771
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,362
; FILING DATE: 21 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-151-771-9

Alignment Scores:
Pred. No.: 30.5 Length: 318
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 10 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-151-771-9 (1-318)
QY 2 TGCACATATTTAGGCGGAATTCGATATTTTCCTTATAGTTATCTATATGCGTA 55
||| |||||: :||| ||||| ||||| |||
Db 94 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyArgVal 111

RESULT 9
US-09-809-391-695
; Sequence 695, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-695

Alignment Scores:
Pred. No.: 30.6 Length: 342
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 11 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-809-391-695 (1-342)
QY 2 TGCACATATTTAGGCGGAATTCGATATTTTCCTTATAGTTATCTATATGCGTA 55
||| |||||: :||| ||||| ||||| |||
Db 77 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyArgVal 94

RESULT 10
US-09-882-171-695
; Sequence 695, Application US/09882171
; Publication No. US20030175858A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11

```



```

Alignment Scores:
Pred. No.: 30.6 Length: 342
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPX_1_63 (1-63) x US-09-882-171-695 (1-342)

```

```

Qy      2  TGCACATATTTTACGCCGAATTCGCATATTTTCTTATAGTATTCCTATATGGTA 55
      |||  |||||:::  ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      77  CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuLeuPheTrgVal 94

RESULT 11
US-10-205-219-131
; Sequence 131, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin

```

```
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 131
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: E124
US-10-205-219-131

Alignment Scores:
Pred. No.: 30.7          Length: 358
Score: 48.00           Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-205-219-131 (1-358)
QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTAGTATTTCTATATGGA 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuLeuPheTyArgVal 110

RESULT 12
US-10-385-450-8
; Sequence 8, Application US/10385450
; Publication No. US20030157683A1
; GENERAL INFORMATION:
; APPLICANT: Lehar, et al., Sophie M.
; TITLE OF INVENTION: APOPTOSIS GENE E124, COMPOSITIONS, AND METHODS OF USE
; FILE REFERENCE: 104322.170DIV
; CURRENT APPLICATION NUMBER: US/10/385,450
; CURRENT FILING DATE: 2003-03-12
; PRIOR FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Murine
US-10-385-450-8

Alignment Scores:
Pred. No.: 30.7          Length: 358
Score: 48.00           Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-385-450-8 (1-358)
QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTAGTATTTCTATATGGA 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuLeuPheTyArgVal 110

RESULT 13
US-10-385-450-9
; Sequence 9, Application US/10385450
; Publication No. US20030157683A1
; GENERAL INFORMATION:
; APPLICANT: Lehar, et al., Sophie M.
```

```
; TITLE OF INVENTION: APOPTOSIS GENE E124, COMPOSITIONS, AND METHODS OF USE
; FILE REFERENCE: 104322.170DIV
; CURRENT APPLICATION NUMBER: US/10/385,450
; CURRENT FILING DATE: 2003-03-12
; PRIOR FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Murine
US-10-385-450-9

Alignment Scores:
Pred. No.: 30.7          Length: 358
Score: 48.00           Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 12 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-385-450-9 (1-358)
QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTAGTATTTCTATATGGA 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuLeuPheTyArgVal 110

RESULT 14
US-10-001-876-176
; Sequence 176, Application US/10001876
; Publication No. US20020177140A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes a
; FILE REFERENCE: DEX-0285
; CURRENT APPLICATION NUMBER: US/10/001,876
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 176
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-876-176

Alignment Scores:
Pred. No.: 40.8          Length: 69
Score: 47.00           Matches: 7
Percent Similarity: 63.16% Conservative: 5
Best Local Similarity: 36.84% Mismatches: 7
Query Match: 40.87% Indels: 0
DB: 14 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-10-001-876-176 (1-69)
QY 4 CACTATTTAGCGGAATTCATATTTTCCTTAGTATTTCTATATGGAAC 60
Db 6 HisLeuPhePheGlnLysCysLeuLeuTyMetIleLeuIleTyTyrSerLysAsn 24

RESULT 15
US-09-939-980-391
; Sequence 391, Application US/09939980
; Patent No. US2002008234A1
; GENERAL INFORMATION:
```

APPLICANT: Black, Michael
Burnham, Martin
Hodgson, John
Knowles, David
Lonetto, Michael
Nicholas, Richard
Pratt, Julie
Reichard, Richard
Rosenberg, Martin
Ward, Judith
TITLE OF INVENTION: NO. US2002082234A1e1 Prokaryotic Polynucleotides,
Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,980
FILING DATE: 27-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,165
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 391:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 391:
US-09-939-980-391
Alignment Scores:
Pred. No.: 88.2 Length: 110
Score: 45.00 Matches: 7
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 0
Query Match: 39.13% Indels: 0
DB: 9 Gaps: 0
US-09-147-052-1_COPY_1_63 (1-63) x US-09-939-980-391 (1-110)
QY 15 GCGGAATTCATATTTTCCTATAGT 41
Db 73 SerGluLeuHisValPheProTyrSer 81

Search completed: October 8, 2003, 19:26:26
JOB time : 5.45109 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:41:47 ; Search time 0.991895 Seconds
(without alignments)
5374.731 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcactatttttagcgaa.....ttctatatgttgcgaactca 63

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-MODEL=frame+n2p.model -DB=us09147052/runat_08102003_154342_29880/app_query.fasta_1.5980
-Q=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154342_29880/app_query.fasta_1.5980
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1_1_126@runat_08102003_154342_29880 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DISPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	865	1	US-07-803-633A-13
2	58	50.4	61	4	US-09-134-001C-4538
3	48	41.7	317	2	US-08-619-362A-8
4	48	41.7	317	2	US-08-790-572-3
5	48	41.7	317	2	US-09-213-398-3
6	48	41.7	318	2	US-08-619-362A-9
7	48	41.7	340	2	US-08-790-572-1
8	48	41.7	340	2	US-09-213-398-1
9	48	41.7	342	4	US-09-149-476-695
10	48	41.7	358	4	US-09-151-771B-8
11	48	41.7	358	4	US-09-151-771B-9
12	48	41.7	359	4	US-09-154-750A-90

C 13	47	45.2	182	2	US-08-874-832-10	Sequence 10, Appl
C 14	47	45.2	182	2	US-08-874-832-11	Sequence 11, Appl
C 15	47	45.2	182	2	US-08-874-832-12	Sequence 12, Appl
C 16	47	45.2	182	2	US-08-874-832-13	Sequence 13, Appl
C 17	47	45.2	182	3	US-09-097-233-10	Sequence 10, Appl
C 18	47	45.2	182	3	US-09-097-233-11	Sequence 11, Appl
C 19	47	45.2	182	3	US-09-097-233-12	Sequence 12, Appl
C 20	47	45.2	182	3	US-09-097-233-13	Sequence 13, Appl
C 21	47	45.2	217	2	US-08-874-832-1	Sequence 1, Appl
C 22	47	45.2	217	2	US-08-874-832-2	Sequence 2, Appl
C 23	47	45.2	217	2	US-08-874-832-3	Sequence 3, Appl
C 24	47	45.2	217	2	US-08-874-832-4	Sequence 4, Appl
C 25	47	45.2	217	3	US-09-097-233-1	Sequence 1, Appl
C 26	47	45.2	217	3	US-09-097-233-2	Sequence 2, Appl
C 27	47	45.2	217	3	US-09-097-233-3	Sequence 3, Appl
C 28	47	45.2	217	3	US-09-097-233-4	Sequence 4, Appl
C 29	47	45.2	614	4	US-09-328-352-6661	Sequence 6661, Ap
C 30	45	39.1	36	2	US-08-598-873-13	Sequence 13, Appl
C 31	45	39.1	36	3	US-08-605-430-13	Sequence 13, Appl
C 32	45	39.1	110	4	US-08-936-165A-391	Sequence 391, App
C 33	45	39.1	456	4	US-09-134-001C-3659	Sequence 3659, Ap
C 34	45	39.1	968	3	US-08-651-999A-7	Sequence 7, Appl
C 35	45	39.1	968	3	US-09-385-752-7	Sequence 7, Appl
C 36	44.5	42.8	780	4	US-09-252-991A-32892	Sequence 32892, A
C 37	44.5	38.7	866	3	US-08-651-999A-1	Sequence 1, Appl
C 38	44.5	38.7	866	3	US-09-385-752-1	Sequence 1, Appl
C 39	43	37.4	76	4	US-09-107-532A-4718	Sequence 4718, Ap
C 40	43	37.4	570	4	US-09-134-001C-2972	Sequence 2972, Ap
C 41	42	36.5	97	4	US-09-663-600A-183	Sequence 183, App
C 42	42	40.4	396	3	US-08-961-083-16	Sequence 16, Appl
C 43	42	40.4	396	4	US-09-536-784-16	Sequence 16, Appl
C 44	42	40.4	600	4	US-08-836-687B-31	Sequence 31, Appl
C 45	42	36.5	633	4	US-09-252-991A-27482	Sequence 27482, A

ALIGNMENTS

RESULT 1
US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025ofu
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 42A
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848

```
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 865 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-803-633A-13

Alignment Scores:
Pred. No.:          5,29e-11          Length:      865
Score:             115.00             Matches:     21
Percent Similarity: 100.00%           Conservative: 0
Best Local Similarity: 100.00%       Mismatches:  0
Query Match:       100.00%           Indels:      0
DB:                1                Gaps:         0

US-09-147-052-1_COPY_1_63 (1-63) x US-07-803-633A-13 (1-865)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTCATATGATGACGAC 60
Db 1 MethIstYrPheA-gAArgAsnCysIlePheLeuIleValIleLeuTyGlyThrAsn 20
QY 61 TCA 63
Db 21 Ser 21

RESULT 2
US-09-134-001C-4538
; Sequence 4538, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4538
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4538

Alignment Scores:
Pred. No.:          0.114             Length:      61
Score:             58.00             Matches:     9
Percent Similarity: 81.25%           Conservative: 4
Best Local Similarity: 56.25%       Mismatches:  3
Query Match:       50.43%           Indels:      0
DB:                4                Gaps:         0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-134-001C-4538 (1-61)
QY 1 ATGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTCATATGATGAC 48
Db 5 LeuHstYrPheAsnArgAsnCysLeuTyLeuIleLeuGluLeu 20

RESULT 3
US-08-619-362A-8
; Sequence 8, Application US/08619362A
; Patent No. 5843659
; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and
; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
; NUMBER OF SEQUENCES: 17
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,362A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 317 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-619-362A-8

Alignment Scores:
Pred. No.:          6.18             Length:      317
Score:             48.00             Matches:     10
Percent Similarity: 66.67%           Conservative: 2
Best Local Similarity: 55.56%       Mismatches:  6
Query Match:       41.74%           Indels:      0
DB:                2                Gaps:         0

US-09-147-052-1_COPY_1_63 (1-63) x US-08-619-362A-8 (1-317)
QY 2 TGCACATTTTAGCGGAATTCATATTTTCTTATAGTATTCATATGATGAC 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyArgVal 110

RESULT 4
US-08-790-572-3
; Sequence 3, Application US/08790572
; Patent No. 5858715
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/790,572
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0204 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 317 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1256606

US-08-790-572-3

Alignment Scores:
Pred. No.: 6.18 Length: 317
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-08-790-572-3 (1-317)

QY 2 TGCATATTTAGGCGGAATTCATATTTTCCTTATAGTATTCTATATGGTA 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyArgVal 110

RESULT 5

US-09-213-398-3
Sequence 3, Application US/09213398
Patent No. 5955429

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,398
FILING DATE: 12/15/98
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/790,572
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0204 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 317 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1256606
US-09-213-398-3

Alignment Scores:
Pred. No.: 6.18 Length: 317
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-213-398-3 (1-317)

QY 2 TGCATATTTAGGCGGAATTCATATTTTCCTTATAGTATTCTATATGGTA 55
Db 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyArgVal 110

RESULT 6

US-08-619-362A-9
Sequence 9, Application US/08619362A
Patent No. 5843659

GENERAL INFORMATION:
APPLICANT: SOPHIE M. LEHAR; and
APPLICANT: BRAYDON C. GUILD
TITLE OF INVENTION: NOVEL APOPTOSIS GENE E124,
TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,362A
FILING DATE: 21 MARCH 1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: WIXON, Henry N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 104322.170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 942-8400
TELEFAX: (202) 942-8484
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-619-362A-9

Alignment Scores:
Pred. No.: 6.18 Length: 318
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-08-619-362A-9 (1-318)

QY 2 TGCATATTTAGGCGGAATTCATATTTTCCTTATAGTATTCTATATGGTA 55

```
Db          94 CysalaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTrpArgVal 111
RESULT 7
US-08-790-572-1
; Sequence 1, Application US/08790572
; Patent No. 5858715
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/790,572
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0204 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 2445356
;
US-08-790-572-1
Alignment Scores:
Pred. No.: 6.23 Length: 340
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-08-790-572-1 (1-340)
QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGGA 55
Db 75 CysalaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTrpArgVal 92
RESULT 8
US-09-213-398-1
; Sequence 1, Application US/09213398
; Patent No. 5955429
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/213,398
; FILING DATE: 12/15/98
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/790,572
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0204 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 2445356
;
US-09-213-398-1
Alignment Scores:
Pred. No.: 6.23 Length: 340
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x US-09-213-398-1 (1-340)
QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGGA 55
Db 75 CysalaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTrpArgVal 92
RESULT 9
US-09-149-476-695
; Sequence 695, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
```



```

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

Alignment Scores:

```

Pred. No.: 6.23 Length: 342
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 4 Gaps: 0

```

```

US-09-147-052-1_COPY_1_63 (1-63) x US-09-149-476-695 (1-342)

```

```

QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTTCTATATGTTATCTATATGTA 55
||| |||||: :||||||| ||||| |||
Db 77 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuPheTrpArgVal 94

```

RESULT 10

```

US-09-151-771B-8
; Sequence 8, Application US/09151771B
; Patent No. 6586204
; GENERAL INFORMATION:
; APPLICANT: Lehman, et al., Sophie M.
; TITLE OF INVENTION: APOPTOSIS GENE E124, COMPOSITIONS, AND METHODS OF USE
; FILE REFERENCE: 104322.170DIV
; CURRENT APPLICATION NUMBER: US/09/151,771B
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Murine
US-09-151-771B-8

```

Alignment Scores:

```

Pred. No.: 6.26 Length: 358
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6

```

```

Query Match: 41.74% Indels: 0
DB: 4 Gaps: 0
US-09-147-052-1_COPY_1_63 (1-63) x US-09-151-771B-8 (1-358)

```

```

QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTTCTATATGTTATCTATATGTA 55
||| |||||: :||||||| ||||| |||
Db 93 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuPheTrpArgVal 110

```

RESULT 11

```

US-09-151-771B-9
; Sequence 9, Application US/09151771B
; Patent No. 6586204
; GENERAL INFORMATION:
; APPLICANT: Lehman, et al., Sophie M.
; TITLE OF INVENTION: APOPTOSIS GENE E124, COMPOSITIONS, AND METHODS OF USE
; FILE REFERENCE: 104322.170DIV
; CURRENT APPLICATION NUMBER: US/09/151,771B
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Murine
US-09-151-771B-9

```

Alignment Scores:

```

Pred. No.: 6.26 Length: 358
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 4 Gaps: 0

```

```

US-09-147-052-1_COPY_1_63 (1-63) x US-09-151-771B-9 (1-358)

```

```

QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTTCTATATGTTATCTATATGTA 55
||| |||||: :||||||| ||||| |||
Db 93 CysAlaTrpAsnGlyValPheTrpPheSerLeuLeuPheTrpArgVal 110

```

RESULT 12

```

US-09-154-750A-90
; Sequence 90, Application US/09154750A
; Patent No. 6432640
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: p53-Induced Apoptosis
; FILE REFERENCE: 1107.75357
; CURRENT APPLICATION NUMBER: US/09/154,750A
; CURRENT FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/059,153
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/079817
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-154-750A-90

```

Alignment Scores:

```

Pred. No.: 6.26 Length: 359
Score: 48.00 Matches: 10
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 4 Gaps: 0

```



```

; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,832
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 2261-1-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-874-832-12

```

```

Alignment Scores:
Pred. No.:      8.54      Length:      182
Score:          47.00     Matches:      8
Percent Similarity: 83.33%  Conservative: 2
Best Local Similarity: 66.67%  Mismatches: 2
Query Match:      45.19%     Indels:      0
DB:               2         Gaps:        0

```

```

US-09-147-052-1_COPY_1_63 (1-63) x US-08-874-832-12 (1-182)

```

```

Qy 39 TATAAGGAATAATGCAATTCGCGCTAATAATAGTG 4
Db 148 TyrIysGluArgLeuGlyLeuProIysIleVal 159

```

```

Search completed: October 8, 2003, 18:20:09
Job time : 1.99189 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:32 ; Search time 2.0777 Seconds
(without alignments)
5832.060 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcactatttttagcgga.....ttctatatgtgtacgaactca 63

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlpl
-O=/cgn2_1/USPTO_spo1/US09147052/runat_08102003_154341_29860/app_query.fasta_1.5980
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052.ecgn_1_1_287/runat_08102003_154341_29860 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	865	1	VGBERB
2	54	47.0	436	2	T20385
3	51	49.0	275	2	F01733
4	51	44.3	352	2	F88957
5	47	45.2	217	2	B26411
6	47	45.2	217	2	A34295
7	47	45.2	217	2	A26411
8	47	45.2	217	2	I45644
9	47	40.9	297	2	E96998
10	46	44.2	305	2	T12744
11	46	40.0	418	2	B72353
12	45	39.1	135	1	G71547
13	45	39.1	135	2	B81702
14	45	39.1	318	2	C81386

15	45	39.1	448	2	G83818	hypothetical prote
16	45	39.1	448	2	G89938	conserved hypothet
17	45	39.1	451	1	E69952	conserved hypothet
c 18	45	43.3	462	1	A56120	steroid hormone re
c 19	45	43.3	462	2	A40716	embryonal long ter
c 20	45	43.3	465	2	A42128	proline-trna ligas
c 21	45	39.1	488	2	A70150	NADH2 dehydrogenas
22	45	39.1	590	1	S34960	polycystic kidney
23	45	39.1	608	2	G02640	hypothetical prote
24	45	39.1	632	2	T24405	hypothetical prote
25	45	39.1	734	2	T24908	hypothetical prote
26	45	39.1	835	2	T47521	respiratory burst
27	44.5	38.7	309	2	T46226	hypothetical prote
c 28	44.5	42.8	754	2	G83025	topoisomerase IV s
c 29	44	38.3	226	2	S41032	hypothetical prote
c 30	44	42.3	265	2	G95058	integrase/recombin
c 31	44	42.3	265	2	G97927	integrase/recombin
32	44	38.3	280	2	AF3650	exopolysaccharide
c 33	44	42.3	573	2	S12838	Ig mu chain precur
34	44	38.3	608	2	B53295	xylanase (EC 3.2.1
c 35	44	38.3	655	1	A54306	proprotein convert
36	44	42.3	3394	2	T18501	hypothetical prote
37	43.5	37.8	300	2	AI0486	dipeptide transpor
38	43.5	37.8	375	2	A81227	conserved hypothet
39	43.5	37.8	375	2	F81999	probable integral
40	43	37.4	59	2	AE2023	hypothetical prote
41	43	37.4	78	2	F81093	hypothetical prote
42	43	37.4	90	2	AC0312	hypothetical prote
43	43	37.4	112	2	B86263	F13k23.19 protein
44	43	37.4	221	2	T52138	eukaryotic cap-bin
45	43	37.4	444	2	E95065	conserved hypothet

ALIGNMENTS

RESULT 1

VGBERB
glycoprotein B precursor - Marek's disease virus (strain RB1B)
C:Species: Marek's disease virus

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A32402; B32402

R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.
J. Gen. Virol. 70, 1789-1804, 1989

A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog

A:Reference number: A32402; MUID:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 <ROS>

A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BA02866.1; PID:g221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271;304-330 <ROS2>

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-865/Product: glycoprotein B #status predicted <GPB>

F:709-728/Domain: transmembrane #status predicted <TM1>

F:732-752/Domain: transmembrane #status predicted <TM2>

F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Alignment Scores:

Pred. No.: 2.02e-11 Length: 865
Score: 115.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x VGBERB (1-865)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTATAGTATTTCTATATATGTTAGTACGAC 60

DB 1 MetHisTyrPheArgAsnCysIlePhePheLeuValIleLeuTyrGlyThrAsn 20

QY 61 TCA 63
|||
Db 21 Ser 21

RESULT 2

T20385
hypothetical protein DH11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T20385
R:Lightning, J.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19265
A:Accession: T20385
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-436 <WIL>
A:Cross-references: EMBL:Z49126; PIDN:CAA88939.1; GSPDB:GNO0020; CESP:DH11.2
A:Experimental source: clone DH11
C:Genetics:
A:Gene: CESP:DH11.2
A:Map position: 2
A:Introns: 41/3; 112/1; 144/1; 320/3; 352/3; 373/1; 406/2
C:Superfamily: Caenorhabditis elegans hypothetical protein DH11.2

Alignment Scores:
Pred. No.: 0.786 Length: 436
Score: 54.00 Matches: 9
Percent Similarity: 68.75% Conservative: 2
Best Local Similarity: 56.25% Mismatches: 5
Query Match: 46.96% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x T20385 (1-436)

QY 7 TATTTAGGCGGAATTCATATTTTCCTTATATGTTATTCATATGCT 54
||||| ||||||||| : : : :
Db 56 TyrPhePheAsnAnCysIlePhePheGlnLeuValAlaTyGly 71

RESULT 3

F90173
conserved hypothetical protein [Imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: F90173
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90173
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <KOR>
A:Cross-references: GB:AE006641; NID:g13813451; PIDN:AAK40645.1; GSPDB:GNO0155
A:Gene: SSO0308
C:Superfamily: hypothetical protein MJ1440

Alignment Scores:
Pred. No.: 2.62 Length: 275
Score: 51.00 Matches: 9
Percent Similarity: 86.67% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 2
Query Match: 49.04% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x F90173 (1-275)

QY 60 GTTCGTACCATATAGATAACTATATAGGAAATAATGCAATCCG 16
: : : : : ||||| ||| : : : : : ||||| : : : : : |||||

Db 52 IleLysThrIleLeuAsnTyrPheLysGluLysTyrSerIlePro 66

RESULT 4

F88957
protein ZK697.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F88957
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: F88957
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <SFO>
A:Cross-references: GB:chr_V; PIDN:AAC48249.1; PID:g2736397; GSPDB:GNO0023; CESP:ZK69
C:Genetics:
A:Gene: ZK697.1
A:Map position: 5

Alignment Scores:
Pred. No.: 2.61 Length: 352
Score: 51.00 Matches: 8
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 0
Query Match: 44.35% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x F88957 (1-352)

QY 19 AATTGCATATTTTCCTTATAGTTATCTATAT 51
||||| ||||||||| : : : : : |||||

Db 91 AspCysIlePhePheLeuLeuLeuLeuLeuTyr 101

RESULT 5

B26411

translation initiation factor eIF-4E - rabbit
N:Alternate names: mRNA cap-binding protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 25-Mar-1988 #sequence_revision 10-Mar-1994 #text_change 26-Aug-1999
C:Accession: S30248; B26411; S16756
R:Rychlik, W.; Rhoads, R.E.
Nucleic Acids Res. 20, 6415, 1992
A:Title: Nucleotide sequence of rabbit eIF-4E cDNA.
A:Reference number: S30248; MUID:93117125; PMID:1475206
A:Accession: S30248
A:Molecule type: mRNA
A:Residues: 1-217 <RYC>
A:Cross-references: EMBL:X61939; NID:g1554; PIDN:CAA43943.1; PID:g1555
R:Rychlik, W.; Domier, L.L.; Gardner, P.R.; Hellmann, G.M.; Rhoads, R.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 945-949, 1987
A:Title: Amino acid sequence of the mRNA cap-binding protein from human tissues.
A:Reference number: A26411; MUID:87147214; PMID:3469651
A:Accession: B26411
A:Molecule type: protein
A:Residues: 8-20; 43-49; 66-88, 'X', 90-91; 96, 'E', 98, 'G', 100-106; 113-117, 'E', 119; 170-172,
C:Superfamily: translation initiation factor eIF-4E

Alignment Scores:
Pred. No.: 13 Length: 217
Score: 47.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 45.19% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x B26411 (1-217)

QY 39 TATAAGGAAATATGCAATTCGCTAAATAGTG 4
||||| : : : : : |||||

Db 183 TyrLysGluArgLeuGlyLeuProProlsIleVal 194

RESULT 6

A34295

translation initiation factor eIF-4E - mouse

C:Species: Mus musculus (house mouse)

C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 23-Feb-1997

C:Accession: A34295

R:Altman, M.; Mueller, P.P.; Pelletier, J.; Sonenberg, N.; Trachsel, H.

J. Biol. Chem. 264, 12145-12147, 1989

A:Title: A mammalian translation initiation factor can substitute for its yeast homolog

A:Reference number: A34295; MUID:89308629; PMID:2663851

A:Accession: A34295

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: mRNA

A:Residues: 1-217 <ALT>

C:Superfamily: translation initiation factor eIF-4E

C:Keywords: phosphoprotein

Alignment Scores:

Pred. No.:	13	Length:	217
Score:	47.00	Matches:	8
Percent Similarity:	83.33%	Conservative:	2
Best Local Similarity:	66.67%	Mismatches:	2
Query Match:	45.19%	Indels:	0
DB:	2	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x A34295 (1-217)

QY 39 TATAAGGAAATATGCAATTCGCGCTAAATAGTG 4

Db 183 TyrLysGluArgLeuGlyLeuProProlsIleVal 194

RESULT 7

A26411

translation initiation factor eIF-4E - human

N:Alternate names: mRNA cap-binding protein

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Aug-1997

R:Rychlik, W.; Domier, L.L.; Gardner, P.R.; Hellmann, G.M.; Rhoads, R.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 945-949, 1987

A:Title: Amino acid sequence of the mRNA cap-binding protein from human tissues.

A:Reference number: A26411; MUID:87147214; PMID:3469651

A:Accession: A26411

A:Molecule type: mRNA

A:Residues: 1-107, 'N', 109-188, 'F', 190-217 <RYC>

A:Note: This sequence has been corrected in reference A38190

R:Rychlik, W.; Domier, L.L.; Gardner, P.R.; Hellmann, G.M.; Rhoads, R.E.

Proc. Natl. Acad. Sci. U.S.A. 89, 1148, 1992

A:Title: Amino acid sequence of the mRNA cap-binding protein from human tissues.

A:Reference number: A38190; MUID:92141219; PMID:1736299

A:Accession: A38190

A:Molecule type: DNA

A:Residues: 107-109; 188-190 <RYC>

A:Note: sequence modified after extraction from NCBI backbone

A:Note: sequence extracted from NCBI backbone (NCBIN:79809, NCBIN:79814)

C:Genetics:

A:Gene: GDB:EIF4E

A:Cross-references: GDB:126371

A:Map position: 18p11.2-18p11.2

C:Superfamily: translation initiation factor eIF-4E

C:Keywords: phosphoprotein

Alignment Scores:

Pred. No.:	13	Length:	217
Score:	47.00	Matches:	8
Percent Similarity:	83.33%	Conservative:	2
Best Local Similarity:	66.67%	Mismatches:	2
Query Match:	45.19%	Indels:	0
DB:	2	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x A26411 (1-217)

QY 39 TATAAGGAAATATGCAATTCGCGCTAAATAGTG 4

Db 183 TyrLysGluArgLeuGlyLeuProProlsIleVal 194

RESULT 8

I49644

translation initiation factor eIF-4E - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-Aug-1999

C:Accession: I49644

R:Jaramillo, M.; Pelletier, J.; Edery, I.; Nielsen, P.J.; Sonenberg, N.

J. Biol. Chem. 266, 10446-10451, 1991

A:Title: Multiple mRNAs encode the murine translation initiation factor eIF-4E.

A:Reference number: I49644; MUID:91344820; PMID:2037592

A:Accession: I49644

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-217 <RES>

A:Cross-references: GB:M61731; NID:g293333; PIDN:AAA37545.1; PID:g293334

C:Genetics:

A:Gene: eIF-4E

C:Superfamily: translation initiation factor eIF-4E

Alignment Scores:

Pred. No.:	13	Length:	217
Score:	47.00	Matches:	8
Percent Similarity:	83.33%	Conservative:	2
Best Local Similarity:	66.67%	Mismatches:	2
Query Match:	45.19%	Indels:	0
DB:	2	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x I49644 (1-217)

QY 39 TATAAGGAAATATGCAATTCGCGCTAAATAGTG 4

Db 183 TyrLysGluArgLeuGlyLeuProProlsIleVal 194

RESULT 9

E96998

4-hydroxybenzoate octaprenyltransferase related protein CAC0800 [imported] - Clostrid

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: E96998

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: E96998

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78776.1; PID:g15023688; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0800

Alignment Scores:

Pred. No.:	13	Length:	297
Score:	47.00	Matches:	6
Percent Similarity:	80.00%	Conservative:	6
Best Local Similarity:	40.00%	Mismatches:	3
Query Match:	40.87%	Indels:	0
DB:	2	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x E96998 (1-297)

QY 10 TTAGCGCGAATTCATTTTCCTATAGTATCTATATGGT 54

Db 111 PheMetAsnAsnValLeuPhePheIleValLeuLeuIleTyrGly 125

RESULT 10

QY	7	TATTTTACGCCGAATTCGATATTTTTCCTATATAGTATTCTATATGTCACGAAC	60
		:::	
Db	183	TyrMetAlaLysValCysIlePheLeuMetIleAspTyrTyrLysAsn	200
RESULT	12		
	G71547		
		Probable protein-disulfide oxidoreductase (EC 1.8.4.-) - Chlamydia trachomatis (serotype	
		C:Species: Chlamydia trachomatis	
		C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999	

G71547
probable protein-disulfide oxidoreductase (EC 1.8.4.-) - Chlamydia trachomatis (serotype C)
C:Species: Chlamydia trachomatis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

CS81386
probable integral membrane protein Cj0421c [imported] - *Campylobacter jejuni* (strain

C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: C81386
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: C81386
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-318 <PAR>
 A:Cross-references: GB:AL139075; GB:AL111168; NID:q6967817; PIDN:CAB74257.1; PID:q696789
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj0421c

Alignment Scores:
 Pred. No.: 28.8 Length: 318
 Score: 45.00 Matches: 8
 Percent Similarity: 75.00% Conservatives: 4
 Best Local Similarity: 50.00% Mismatches: 4
 Query Match: 39.13% Indels: 0
 DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x C81386 (1-318)

QY 7 TATTTAGCGAATGCATATTTTCCTATAGTATTCATATGTT 54
 Db 24 TyrPheLysAsnSerLysIlePhePheLeuThrLeuValLeuLeuGly 39

RESULT 15

G83818

Hypothetical protein BH1351 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: G83818

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83818

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-448 <STO>

A:Cross-references: GB:AF001511; GB:BA000004; NID:gi0173727; PIDN:BA05070.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1351

C:Superfamily: conserved hypothetical protein b0835

Alignment Scores:
 Pred. No.: 28.6 Length: 448
 Score: 45.00 Matches: 7
 Percent Similarity: 100.00% Conservatives: 2
 Best Local Similarity: 77.78% Mismatches: 0
 Query Match: 39.13% Indels: 0
 DB: 2 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x G83818 (1-448)

QY 15 CGGGAATGCATATTTTCCTATAGT 41

Db 324 SerGluLeuHisValPheProTyrSer 332

Search completed: October 8, 2003, 18:17:16
 Job time : 5.0777 secs

THIS PAGE BLANK (08970)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:40:02 : Search time 119.423 Seconds
(without alignments)
5924.998 Million cell updates/sec

Title: US-09-147-052-1

Perfect score: 2326

Sequence: 1 atgactattttaggcgga.....gacccgggtacatttttataa 1371

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -US09147052/runat_08102003_154341_29846/app_query.fasta_1.5980
-O=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154341_29846/app_query.fasta_1.5980
-DB=SPTRMBL_23 -OFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=pcg -NORM=ext -HEAPSIZE=500 -WINLEN=0 -MAXLEN=2000000000
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	1200	51.6	271 2	Q49464 mycoplasma

Alignment Scores:

RESULT 1	ID	Q49464	PRELIMINARY;	PRT;	271 AA.
AC	Q49464;				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)			
DT	01-MAY-2000	(TrEMBLrel. 13, Last annotation update)			
DE	TN-1	(Fragment)			
GN	TN-1.				
OS	Mycoplasma gallisepticum.				
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.				
OX	NCBI_TaxID=2096;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94025893; PubMed=8212828;				
RA	Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,				
RA	Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.;				
RT	"Cloning and DNA sequence of a 29 kilodalton polypeptide gene of				
RT	Mycoplasma gallisepticum as a possible protective antigen."				
RL	Vaccine 11:1061-1066(1993).				
DR	EMBL; S65869; AAB28343.2; -				
FT	NON_TER 271				
SQ	SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;				

2	845	36.3	671	2	Q918D6
3	832	35.8	702	2	Q49499
4	829	35.6	632	2	Q9XCG8
5	805	34.6	584	2	Q9KH13
6	761	32.7	680	2	Q9KH14
7	754	32.4	702	2	Q49498
8	750.5	32.3	645	2	Q9ZID1
9	749.5	32.2	644	2	Q9ZHR9
10	749	32.2	650	2	Q49495
11	745.5	32.1	649	2	Q49497
12	740.5	31.8	486	2	Q8RLX9
13	715.5	30.8	647	2	Q49468
14	704	30.3	656	2	Q9KH15
15	700.5	30.1	703	2	O05122
16	619.5	26.6	419	2	Q918D5
17	578	24.8	386	2	Q49500
18	557.5	24.0	367	2	Q9XCG7
19	314	13.5	805	12	Q98V44
20	314	13.5	805	12	Q98V45
21	314	13.5	865	12	Q83291
22	314	13.5	865	12	Q8JLW2
23	314	13.5	865	12	Q8JLW3
24	314	13.5	865	12	Q8JLW4
25	196.5	8.4	864	12	Q69408
26	196.5	8.4	870	12	Q9DPQ9
27	196.5	8.4	870	12	Q9E1G4
28	189.5	8.1	320	2	Q49496
29	185.5	8.0	9439	16	Q8CP76
30	177.5	7.6	865	12	Q69406
31	177.5	7.6	865	12	Q9PWZ1
32	177	7.6	6713	16	Q99054
33	177	7.6	6713	16	Q931R6
34	173.5	7.5	933	2	Q53653
35	173	7.4	946	16	Q8NXJ1
36	173	7.4	1302	2	Q49547
37	170	7.3	9904	16	Q8NW06
38	168.5	7.2	810	5	Q814Y6
39	166.5	7.2	3890	16	Q99053
40	166.5	7.2	4688	16	Q9PQ08
41	166	7.1	661	16	Q8YMJ8
42	165.5	7.1	2481	16	Q99QR6
43	164.5	7.1	721	16	Q8XH94
44	164.5	7.1	1774	5	Q81AU8
45	163.5	7.0	2462	16	Q8RGZ3

ALIGNMENTS

```

Pred. No.: 3.3e-59 Length: 271
Score: 1200.00 Matches: 238
Percent Similarity: 98.37% Conservative: 3
Best Local Similarity: 97.14% Mismatches: 4
Query Match: 51.59% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1 (1-1371) x Q49464 (1-271)
QY 190 TGTATGCTTACTATAAAGATGCAACCAATAATGCGCAACCAATTAAGAACA 249
DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
QY 250 GCGGAATGGAGTAAACAGATCTAATCAATGCTAAACGGATGACATTAAGTTCATCAAA 309
DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
QY 310 GACTATGCCAAGATGAAGCTAGTATTCATCTCTGTTATAGTGAAGCTGAACAGTAAAC 369
DB 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
QY 370 AATAACCTTAATGCAACATTAACAACATAAATGCGTAAACATAATTAAGATCAGCC 429
DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGGATAAAGACACTTTTCATATGATGAACACCAAAATTAAGTCAA 489
DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
QY 490 GCATACAAAGCACTAAAACACCTTTAGAACAACGCTGCTACTAACTTGAAGGTTGTCA 549
DB 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
QY 550 TCACTGCTTATATCAAAATCGCAATATTAAGTGGATCATACAATAAGCTAGTAGT 609
DB 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
QY 610 TTAATACTAAACACTAGATCCACTAAATGCGGGAACGCTTTTAGATTTCTAATGAGATT 669
DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
QY 670 ACTACAGCTAATAAGATATTAATATACGTTATCAACTATTAATGAACAAAAGACTAAT 729
DB 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GCTGATGCATATCTAATAGTTTATTAATAAAGTGATTCAAATATGAAACAAAGTTT 789
DB 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226
QY 790 GTAGGACTTTTACAAACGCTAATGTTCACACCTTCAACACTACAGTTTGTGCTTTTGT 849
DB 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
QY 850 GCTGATGTACACCGCTCAATTAATTAATATGCAAGAAGGACCGTTTGGAAATGGTGATGA 909
DB 247 AlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValThrPheAsnGlyAspGlu 266
QY 910 CCITCAAGTAGAATT 924
DB 267 ProSerSerArgIle 271

RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
ID Q9L8D6;
AC Q9L8D6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PKGA-like protein 9.2.
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]

```

```

SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RT May J.D., Hughlett M.B.;
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAF29524.1;
SQ SEQUENCE 671 AA; 71898 MW; 4D6AB6B59175D679 CRC64;

Alignment Scores:
Pred. No.: 1.65e-39 Length: 671
Score: 845.00 Matches: 186
Percent Similarity: 58.19% Conservative: 77
Best Local Similarity: 41.15% Mismatches: 127
Query Match: 36.33% Indels: 62
DB: 2 Gaps: 10

US-09-147-052-1 (1-1371) x Q9L8D6 (1-671)
QY 157 GTGGTTCAACCGTGATCCGCTCTAGAATTCGGCTGTATGCTATTACTATAAAGATGCA 216
DB 15 IleGlySerPheValMetLeuAlaAlaSerCysThrSerAlaThrThrProThrPro 34
QY 217 AACCCA-----AATAATGCG 231
DB 35 AsnProGluProLysProAspProMetProAsnProProSerGlyAspMetAsnGlyGly 54
QY 232 CAACAC-----CAATAGACGACG 252
DB 55 AspThrAsnProGlyAsnAspGlyGlyMetGluAsnSerAlaGlnGlnLeuAlaAla 74
QY 253 CGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGCA 312
DB 75 LysLysGluLeuSerAspLeuLeuAlaThrGlnSerSerAsnLeuAlaLysTyrAlaAsp 94
QY 313 TATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTTAACAAT 372
DB 95 TyrThrAsnIleGlnAsnThrLeuThrAlaAlaTyrThrAlaLysSerThrSerAsp 114
QY 373 AACCTTAATGCAACATTAGAACAACTAAATGCTAAACTAATTTAGATCAGCCATC 432
DB 115 AsnThrSerValThrLeuGluGlnValLysSerAlaThrSerThrLeuGlnAlaAla 134
QY 433 AACCAAGCTAATACGGATAAACAGCTTTTGTATAATGAACACCCCAATTTAGTTGAAGCA 492
DB 135 AspThrAlaAlaSerSerLysThrSerPheAspGluLysAsnProGluLeuLysAla 154
QY 493 TACAAAGCCTAAACACCACTTTAGAACACACGCTGCTACTAACCTTGAAGGTTGCTCATCA 552
DB 155 TyrTyrAlaLeuLysGluThrLeuLysAsnGluGluThrValLeuSerGlyLeuThrAsp 174
QY 553 ACTGCTTATATCAAAATTCGCAATAATTTAGTGGATCTATACATAAAGCTAGTTTGA 612
DB 175 SerAsnPheAlaThrIleLysThrAsnLeuThrAlaLeuTyrGlnSerGlyLysAspPhe 194
QY 613 ATAACTAAACACTAGATCCACTAAATTAATGCTTATCAACTATTATGAACAAAAGCTAATGCT 672
DB 195 ValLysAlaThrLeuAspProValSerGlyAsnAla--ProGlnIleAlaAspIleThr 213
QY 673 ACAGCTAATAAGAATATTATTAATACCTTATCACTATTATTAATGAACAAAAGCTAATGCT 732
DB 214 LysAlaAspLysAspIleAlaAspAlaValSerLysLeuGluThrTrpLysThrAsnAla 233
QY 733 GATGCAATTATCTAATAGTTTATTAAAAAGTGTATCAAAATGAATGAACAAAGTTTGA 792
DB 234 AsnThrLeuAlaThrSerPheValLysGluValLeuValLysAsnLysLeuThrGlyLe 253
QY 793 GGGACTTTTACAACGCTAATGTTCAACCTTCAACCTTCAACCTTGTGCTTTGTAGTGC 852
DB 254 AspThr---ThrAsnAsnArgGluGlnProGlyAsnTyrSerPheValGlyTyrSerVal 272
QY 853 GATGTAACACCC-----GTCANTTATAATATGCAAGACGACCGTTTGGAAAT 900

```


265	ArgAlaHisAsnGlnGluGlnProAlaAsnTrpSerPheAlaGlyTrpSerValAspLeu	284
859	ACACCC-----GTCAATTATAAAATATGTCGAAAGACGACC	891
285	ThrThrGlySerThrGlyAsnSerGlnAsnLeuProAsnTrpAsnPheAlaGlnArgLys	304
892	TTTGGGAATGGTGAT-----GAACCTTCAAGTAGAATCTCTGCGAAACACGAATAGT	942
305	ValTrpThrSerGluGlyGlnGlnThrGlyLysThrAlaLeuValSerSerProValSer	324
943	ATCACAGATGTTCTTGGGATTTATAGTTTACGTGGGAACAAACACGGAAGTACCAATTTAGT	1002
325	AlaThrAspValSerTrpIleTrpSerLeuAlaGlyGluGlyThrLysTrpThrLeuSer	344
1003	TTTAGCAACATATGTCATCACTGGGTATTTATATTTATTTCCCTATATAGTTGGTTAAAGCA	1062
345	PheGluTrpTyrglyProAspThrAlaPheLeuTrpPheProTrpLysLeuValLysGln	364
1063	GCTGATGCTAATAACGTTGGATTACATACAAATTAATAATGCGAAATGTTCAACAAGTT	1122
365	AlaAspSerSerValAlaLeuGlnTyrSerLeuAsnLysThrSerSerLysLeuIle	384
1123	GAGTTT-----	1128
385	AsnPheGluProAlaLysThrMetProThrAsnAlaAspGlnSerGluAsnGlyValAla	404
1129	GCACACTCAACT-----	1143
405	ThrThrSerThrThrGluGlyArgSerSerSerGluValLeuValAlaAspGluValAla	424
1144	GCAAYATAACTACAGCTAATCCAACTCCACAGCTTGATGAGATTTAAAGTTGCTAAATC	1203
425	AlaValAsnAsnGluMetAsnProThrProThrValSerAspIleAsnIleAlaLysVal	444
1204	GTTTTATCAGGTTTAGATTTGGCCAAACACAACTCGAATAGTGTTCACACGGGTGAA	1263
445	ThrLeuSerGlyLeuThrPheGlyGluAsnThrIleGluPheSerValProThr-----	462
1264	GGAATATGATTAAGTTGGCCCAATGATTGGCAACATTTATCTTAGCTCAAAATGAAAT	1323
463	-----AsnLysValAlaProMetIleGlyAsnMetTrpLeuThrSerAsnSerGly	479
1324	AATGCTGATAGATC	1338
480	SerGlnGlyLysIle	484
RESULT 6		
ID	Q9KH14	PRELIMINARY; PRT; 680 AA.
AC	Q9KH14;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)	
DE	Adhesin pmGAL.3.	
GN	pmGAL.3.	
OS	Mycoplasma gallisepticum.	
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.	
OX	NCBI_TaxID=2096;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=HS;	
RT	Shen Q.C., Bi D.R., Weng C.J.;	
RT	"Sequence analysis of the pmGA multigene family of Mycoplasma	
RT	gallisepticum strain HS."	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF275312; AAF91414.1; -	
SQ	SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFA7 CRC64;	
Alignment Scores:		
Pred. No.:	7.68e-35	Length: 680
Score:	761.00	Matches: 164
Percent Similarity:	75.04%	Conservative: 71
Best Local Similarity:	39.81%	Mismatches: 129


```

Db 169 LysThrThrLeuLysSerLysThrThrSerLeuGluGlyLeuSerGluAsnLysThrGly 188
Qy 565 CAATTCGCAATAATTTAGTGGATCTATACATAAGCTAGTGTAAATAACTAAACA 624
Db 189 GlyTleLysAsnHisLeuSerLysLeuPheAspThrGlySerAlaIleThrAlaLysThr 208
Qy 625 CTAGATCCACTAAATGGG-----GGAACGGCTTTTAGATCTTAATGAGATTACTACACT 678
Db 209 LeuAspProThrSerGlyGluArgProThrLeuGluLysValAsnGlu-----Ala 225
Qy 679 AATAAGAAATATAATAATACCTTATCA-----ACTATTAAATGAACAAAGACTAATGCT 732
Db 226 AsnAsnGlyIleLysMetAlaIleSerProGluSerLeuLysLysThrLysGlyAsnAla 245
Qy 733 GATGATTAATCATAGTTTATTAAAGAGTATCAAAATATGACAAAGTTTGTGA 792
Db 246 AspLysPhe---AsnGluPheGluLysAsnProLeuSerLysGluLysLeuLysSerThr 264
Qy 793 GGGACTTTTACAAACGGCTAATGTTCAACCTTCAACTACAGTCTTTGCTTTTAGTCT 852
Db 265 SerAspThrAlaHisAsnGlnGluGlnProAlaAsnTrpSerPheAlaIleThrSerVal 284
Qy 853 GATGTAACACCC-----GTCAATTATAATATGCAAGAGGACCGTTGG 897
Db 285 AspLeuThrSerAsnSerGlnAsnLeuProAsnTrpAsnPheAlaGlnArgLysValTrp 304
Qy 898 AATGGTGAT-----GAACCTTCAAGTAGAATCTTCGCAACAGCAAT---AGTATCACA 948
Db 305 ThrSerGluAsnGlnProGlyLysThrAlaLeuValSerProValSerAlaThr 324
Qy 949 GATGTTCTTCGATTTATAGTTTACGCGAACAACAGCAAGTACCAATTTAGTTTACG 1008
Db 325 AspValSerTrpIleThrSerLeuAlaGlyGluGlyThrLysThrLeuThrPheGlu 344
Qy 1009 AACTATGTCATCAACGGTGTATTATATATTCCTTATAGTTGGTTAAAGCAGCTGAT 1068
Db 345 TyrThrGlyProAspAsnAlaPheLeuThrLeuProThrLysLeuValLysAlaAlaAsp 364
Qy 1069 GCTAATACGTTGGATTACATACAAATTAATGGAATGTCACAAAGTTGAGTTT 1128
Db 365 SerSerValAlaLeuGlnThrSerLeuAsnLysThrSerSerLysLeuIleAsnPhe 384
Qy 1129 -----GCCACT 1134
Db 385 LysProAlaGluThrValSerThrAsnThrAspGlnSerGluAsnGluValAlaThrThr 404
Qy 1135 TCAACT-----AGTGCAAT 1149
Db 405 SerThrThrGluAlaArgSerSerThrLysValLeuValAlaAspGluAlaAlaThrSer 424
Qy 1150 AATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTA 1209
Db 425 AsnAsnGluMetAsnHisThrProThrValSerAspIleAsnIleAlaLysValThrLeu 444
Qy 1210 TCAGTTTAAGATTGGCCAAACACACATCGAATTAAGTGTCCACCGGGTGAAGGAAT 1269
Db 445 SerGlyLeuThrPheGlyGluAsnThrIleGluPheSerValProGluGly----- 461
Qy 1270 ATGAATAAGTTGGCCCAATCATTTGGCAACATTTATCTTACTCAATGAAATAATGCT 1329
Db 462 -----LysValAlaProMetIleGlyAsnMetThrLeuThrSerAsnSerGluSerGln 479
Qy 1330 GATAAGATC 1338
Db 480 Vallystle 482

```

RESULT 8

```

Q92ID1 ID Q92ID1 PRELIMINARY: PRT: 645 AA.
AC Q92ID1.
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

```

```

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE M9 protein.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pMCA
RT family.";
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAB055 CRC64;

Alignment Scores:
Pred. No.: 2,96e-34 Length: 645
Score: 750.50 Matches: 177
Percent Similarity: 55.96% Conservative: 67
Best Local Similarity: 40.60% Mismatches: 133
Query Match: 32.27% Indels: 59
DB: 2 Gaps: 12

US-09-147-052-1 (1-1371) x Q92ID1 (1-645)
Qy 157 GTGGTTCACCGTGATCGCTAGAAATTCGGCTGTATGCTTATTACTAAAAA----- 210
Db 15 IlegSerPheValMetLeuAlaAlaSerCysThrThrProThrProAsnProAsn 34
Qy 211 -----GATGCAAAACCCAAATAATGGCAA----- 234
Db 35 ProProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGlyGlnGlyMetMet 54
Qy 235 -----ACCAATTTAGAACCGCGGGAATGAGTTAAACAGATCTTAATCAATGCT 282
Db 55 AsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThrValPheAspSer 74
Qy 283 AAAGCGATGACATTAAGTTCTACTACAAGCATGTCGAAGATTGAAGTAGTTTATCATCT 342
Db 75 LysAlaLysAsnLeuGlyLeuThrValAspThrLysLysThrGlnAspThrLeuThrLys 94
Qy 343 GCTTATAGTGAAGCTGAACAGTTAAACCTTAATCACTTAATGCAACATTAGAACCACTAAA 402
Db 95 AlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThrGlnAsnLeuAsn 114
Qy 403 ATGCTAAACTAATTTAGATTCACCATCAACCAAGCTAATACGGATAAAACGACTTTT 462
Db 115 GluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSerLysGlnThrPhe 134
Qy 463 GATATGAACCCCAATTTAGTTGAAGCATACAAAGCACATAAACCACCTTTAGAACAA 522
Db 135 AspGluGlnHisAlaGluLeuValLysValThrGluGluLeuLysThrThrLeuSerAsn 154
Qy 523 CGTGCTACTAACTTGAGTTTGTCACTCACTGCTTATAATCAATTCGCAATAATTTA 582
Db 155 GluThrAlaThrLeuAlaProThrAlaAlaGlnThrAlaGlyIleLysMetHisLeu 174
Qy 583 GTGATCTATACATAAAGCTAGTAGTTTAACTAAACACTAGATCCACTAAATGGG 642
Db 175 SerGlyLeuThrAspAlaGlyLysAlaIleThrThrLysThrLeuGluProValGluGly 194
Qy 643 GGAACGCTTTTAGATTCCTAATGAGATTACTACAGCTAATAAGATATTAATAATACGT 702
Db 195 AspPro-----LeuThrAlaAspValValMetMetAlaAsnThrLysIleValGluAlaIle 213
Qy 703 -----TCAACTATTAAAGCAAAAGACTAATGCTGATGCTTATCTAATGCTTTTAT 756
Db 214 LysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAlaAspSerPheVal 233
Qy 757 AAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTT 816

```

```

|||||:|||||:  ...  |||  |||
Db 234 LysGlnValLeuValLysGluLysIleThrGlyValGluGluAlaHisAsn---LysAla 252
QY 817 CAACCTCAAACTACAGTTTGTCTTGTAGTCTGCTGATGTAACACCCGTC----- 867
Db 253 GlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThrGlyThrValThrGly 272
QY 868 -----AATTATAATATCCAGAGGAGGACCGTTGG-----AATGGTGAGAA 909
Db 273 GlnThrSerIleProAsnThrPaspTyrAlaGlnArgThrIlePheThrAsnGlyAspGlu 292
QY 910 CCTTCAAGTAGAATCTTGCAACACG-----AATAGATAC 945
Db 293 Pro-----ArgSerIleSerAsnThrProAlaAspGlyGlnThrMetValGlnProLeu 310
QY 946 ACAGATCTTCTTGAGATTATAGTTAGCTGGGAACAAACAGAGTACCAGTTAGTTT 1005
Db 311 SerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLysTyrThrLeuGluPhe 330
QY 1006 AGCAACTATGCTCATCACTGGTTATTTATATTCCTTATAGTTGGTTAAAGCAGCT 1065
Db 331 ThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLysLeuValAsnThrSer 350
QY 1066 GATGCTAATACGTTGGATTACATAACAAATTAATGAAATGTTCAACAAGTTGAG 1125
Db 351 AspGlnMetLysLeuGlyLeuGluTyrLysLeuAsnAsp----- 363
QY 1126 TTGGCCACTCAACTAGTGCA-----AATATACTACAGCTAATCCAACT 1170
Db 364 ---AlaThrGluProSerAlaIleThrPheGlySerGluGlnThrMetAsnGlyLysThr 382
QY 1171 CCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTTTAAAGTTGGCCAA 1230
Db 383 ProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeuLysPheGlySer 402
QY 1231 AACCAATCAATAGTGTCCACGCGGTGAAGGAAATATGAATGAAGTTGCGCAATG 1290
Db 403 AsnLysIleGluPheSerValProAla-----GluLysValSerProMet 417
QY 1291 ATTGGCAACATTATCTAGCTCAATGAAATAATGCTGATAGATC 1338
Db 418 IleGlyAsnMetTyrLeuSerSerProAsnAsnTrpAsnLysIle 433

```

RESULT 9

```

Q92HR9 ID Q92HR9 PRELIMINARY; PRT; 644 AA.
AC Q92HR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Hypothetical 69.9 kDa protein.
OC Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EG31;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmGA
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

```

Alignment Scores:

```

Pred. No.: 3,36e-34 Length: 644
Score: 749.50 Matches: 177
Percent Similarity: 56.19% Conservative: 68
Best Local Similarity: 40.60% Mismatches: 132
Query Match: 32.22% Indels: 59

```

```

Db: 2 Gaps: 12
US-09-147-052-1 (1-1371) x Q92HR9 (1-644)
QY 157 GTGGTTTCAACCGTGATCCGCTCTAGATTCGGCTGTATGTCTATTACTATAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProAsnProAsn 34
QY 211 -----GATGCAAAACCCAAATTAATGCGCAA----- 234
Db 35 ProProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGlyGlyMetMet 54
QY 235 -----ACCAATTAGACGCGAGTAACTAGTAACTAGATCTAATCAATCAATGCT 282
Db 55 AsnAlaAlaSerGlnGlnLeuAlaAlaAlaArgMetGlyLeuThrThrValPheAspSer 74
QY 283 AAGCCGATGACATTACCTTCACAGACTATGCGCAAGACTGGAAGCTAGTTATCATCT 342
Db 75 LysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGlnAspThrLeuThrLys 94
QY 343 GCTTATAGTGAAGCTGAAACAGTTAACTAATACCTTAATGCAACATTAGCAACAACTAAA 402
Db 95 AlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerThrThrGlnAsnLeuAsn 114
QY 403 ATGGCTAAACCTAATTTAGATCAGCCATCAACAGCTAATACGATAAACAGCTTTT 462
Db 115 GluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSerLysGlnThrPhe 134
QY 463 GATAATGAACCCCAAAATTTAGTTGAGCATCAAGCATCAAGCACTAAAACCACTTTTAAACNA 522
Db 135 AspGluGlnHisAlaGluLeuValLysValTyrGluGluLysThrThrLeuSerAsn 154
QY 523 CGTGCTACTAATCACTGAAGTTTGTCACTCACTGCTTATATCAATCAATTCGCAATATTA 582
Db 155 GluThrAlaThrLeuAlaProTyrAlaAlaAlaGlnTyrAlaGlyLysMetHisLeu 174
QY 583 GTGGATCTATACAATAAAGCTAGTAGTTAATACTAAACACTAGATCCACTAAATGGG 642
Db 175 SerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLysLeuGluProValGluGly 194
QY 643 GGAACGCTTTAGATTCTAATGAGATTACTACAGCTAATGAATATTAATTAATACGTTA 702
Db 195 AspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIleValGluAlaIle 213
QY 703 -----TCACTATTAAAGCAAAAGACTAATGCTGATGCTATCTATCTAATAGTTTATT 756
Db 214 LysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAlaAspSerPheVal 233
QY 757 AAAAAGTGATTCAAAATAAATGAACAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTT 816
Db 234 LysGlnValLeuValLysGluLysIleThrGlyValGluGluAlaHisAsnLysSer--- 252
QY 817 CAACCTTCAAACTACAGTTTGTGCTTTTGTGCTGATGCTGATGCTAATCTAATAGTTTATT 867
Db 253 GlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThrGlyThrAlaAsnGly 272
QY 868 -----AATTATAATATGCAAGAGGACCGTTGG-----AATGGTGATGAA 909
Db 273 GlnThrSerIleProAsnTrpAsnTyrAlaGlnArgThrIlePheThrAsnGlyAspGlu 292
QY 910 CCTTCAAGTAGAATCTTGCAACACG-----AATAGATAC 945
Db 293 Pro-----ArgSerValSerAsnThrProValAspGlyGlnThrMetAlaGlnProLeu 310
QY 946 ACAGATGTTTCTTGAGATTATAGTTTGTGCTGGAACAAACAGAGTACCAGTTAGTTT 1005
Db 311 SerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLysTyrThrLeuGluPhe 330
QY 1006 AGCAACTATGCTCATCACTGGTTATTTATATTCCTTATAGTTGGTTAAAGCAGCT 1065
Db 331 ThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLysLeuValAsnThrSer 350
QY 1066 GATGCTAATACGTTGGATTACATAACAAATTAATGAAATGTTCAACAAGTTGAG 1125

```

```

Db      351  |||  :::::::::::::::::::::::::::: 363
QY      1126  TTTGGCAGTTCAACTAGTGCA-----AATAFACCTACAGCTAATCCAAC 1170
Db      364  ---AlaThrLysProSerAlaIleThrPheGlySerGluGlnThrMetAsnGlyLysThr 382
QY      1171  CCAGCAGTGTAGTAAAGTTGCTAAATCGTTTATCATAGGTTTAAGATTTCGCCAA 1230
Db      383  ProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeuAsnPheGlySer 402
QY      1231  AACACAATCGAATTAAAGTTCCCAACGGGTGAAGCAATATCAATAAAGTTTCGCCCAATG 1290
Db      403  AsnLysIleGluPheSerValPro-----ValGluLysValSerPromet 417
QY      1291  AFTGCAACATTTATCTTAGCTCAATGAAATAATGCTGATAAGATC 1338
Db      418  IleGlyAsnMetTyrLeuSerSerProAsnAsnTrpAsnLysIle 433

RESULT 10
Q49495 PRELIMINARY; PRT; 650 AA.
AC Q49495; O08060;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Haemagglutinin precursor.
GN PMGAL.1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0714; AAB50152.1; -.
KW SIGNAL.
FT SIGNAL. 1 25 POTENTIAL.
FT CHAIN 26 650 HAEMAGGLUTININ.
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBB CRC64;

Alignment Scores:
Pred. No.: 3 58e-34 Length: 650
Score: 749.00 Matches: 178
Percent Similarity: 55.61% Conservative: 65
Best Local Similarity: 40.73% Mismatches: 138
Query Match: 32.20% Indels: 56
DB: 2 Gaps: 12

US-09-147-052-1 (1-1371) x Q49495 (1-650)
QY 157 GTGGCTTCAACCGTATCGCTAGAAATTCGGCTGTATGCTATTACTATAAAGATGCA 216
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProSerProAla 34
QY 217 -----AACCAAAATAT 228
Db 35 ProAsnProAsnProProSerAsnGlyGlyMetAsnGlyGlyAsnIleAsnProGlyAsp 54
QY 229 GSCCAA-----ACCAATAGAACGCGCGAATGAGGTAAACA 267
Db 55 GlyGlnGlyMetMetAsnAlaAlaGlnGluLeuAlaAlaAlaArgMetGlyLeuThr 74

```

```

QY 268 GATCTAATAACATGCTAAACCGATGACATTAGCTTCACTACAAGACTATATCCCAAGATTGAA 327
Db 75 ThrValPheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysThrGln 94
QY 328 GCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGATTAAACAATAACCTTAATGCAACA 387
Db 95 AsnThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerThr 114
QY 388 TTAGAACAACTAAATAAGCTTAAACTAATTTAGATCAGCCATCAACCAAGCTATACG 447
Db 115 ThrGlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThr 134
QY 448 GATAAACAGCTTTTGTATATCAACACCAAAATTTAGTTGAAGCATATCAACAGCACTAAA 507
Db 135 SerLysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLys 154
QY 508 ACCACTTTAGAACAAACGCTCTACTAACCCTTGAAGGTTTGTCACTCAACTGCTTATATCAA 567
Db 155 ThrThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGly 174
QY 568 ATTCGCAATAATTTAGTGGATCTATACAATAAGCTAGTAGTTTAAATACTAAAACACTA 627
Db 175 IleLysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeu 194
QY 628 GATCCACTAAATGGGGACGCTTTTAGATTCTTAATGAGATTACTACACTTAATAAGAAAT 587
Db 195 GluProValGluGlyAspPro---LeuThrAlaGlyAlaValThrMetAlaAsnThrLys 213
QY 688 ATTAATAATACGTTA-----TCAACTATTATGAACAAAGACTAATCTGATGCAATTA 741
Db 214 IleValGluAlaIleLysAspGluValLeuAsnProLysLysGluAsnAlaThrLysLeu 233
QY 742 TCTAATAGTTTATTAAAAAGTGATCAAAATAATGAACAAAGTTTGTAGGAGCTTTT 801
Db 234 AlaAspSerPheValLysGlnValLeuValLysGluLysIleThrGlyValGluAla 253
QY 802 ACAAAACGCTAATGTTCAACCTTCAACTACACTACAGTTTGTGCTTTAGTCTGATGTAACA 861
Db 254 HisAsn---LysAlaGlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThr 272
QY 862 CCCGTC-----AATTATAATATGCAAGAGGACCGTTTCG 897
Db 273 GlyThrAlaAsnGlyGlnThrSerIleProAsnTrpAsnTyrAlaGlnArgThrIlePhe 292
QY 898 ---AATGCTGATGAACCTTCAAGTAGAATCTTGCACAAACAG----- 936
Db 293 ThrAsnGlyAspGluPro-----ArgSerValSerAsnThrProValAspGlyGlnThr 310
QY 937 -----AATAGTATCACAGATGTTTCTTGGATTTATAGTTTACGTGGAACAACACGAG 990
Db 311 MetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLys 330
QY 991 TACCAATTTAGTTTGTAGCAACTATGTCATCACTCACTGTTTATTATTATTCCTTATAAG 1050
Db 331 TyrThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 350
QY 1051 TTGGTTAAAGCAGCTGCTAATAACGTTGGATTACATAACAATAATAATGCAAT 1110
Db 351 LeuValAsnThrSerAspGlnMetLysLeuGlyLeuGluTyrLysLeuAsnAspAla--- 369
QY 1111 GTTCAACAAGTTGAGTTTGGCCACTTCACTAGTGCATAATAACTACTACACTAAT---CCA 1167
Db 370 -----ThrGluProSerAlaIleThrPheGlyAsnGluGlnThrMetAsnGlyLys 386
QY 1168 ACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATTCAGGTTTAAGATTGCG 1227
Db 387 ThrProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeuIlePheGly 406
QY 1228 CAAAACACATCGAATTAAGTTTCCAGCGGTGAAGGAAATATGAATAAAGTTGCGCCA 1287
Db 407 SerAsnLysIleGluPheSerValProAla-----GluLysValSerPro 421
QY 1288 ATGATTGGCAACATTTATCTTTAGCTCAATGAAATAATGCTGATAAGATC 1338

```


DR EMBL; AY065985; AAL58980.1; --
FT NON_TER 486 486
SQ SEQUENCE 486 AA; 51844 MW; 94912DD7A09FB911 CRC64;

Alignment Scores:

Pred. No.: 1,09e-33 Length: 486
Score: 740.50 Matches: 174
Percent Similarity: 52.83% Conservative: 69
Best Local Similarity: 37.83% Mismatches: 128
Query Match: 31.84% Indels: 89
DB: 2 Gaps: 12

US-09-147-052-1 (1-1371) x Q9RLX9 (1-486)

```
QY 151 CCACCAGTGGTTCACCGTGATCGCTAGATTGGCTGTATCTATTACTAAAAA 210
DB 152 ProProSerGlyGly-----AsnMetAsnGlyGly 61
QY 211 GATGCAAAACCCAAATPAATGGC-----CAAAACCAATTAGAACACGC 252
DB 62 AspThrAsnProGlyAsnGlyGlyMetAspAsnAlaValGlnGlnLeuAlaAla 81
QY 253 CGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGAC 312
DB 82 LysThrAlaLeuThrThrLeuLeuAsnGlyGlnThrGluLysValGlyLeuTyrAsnAsp 101
QY 313 TATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTAAACAT 372
DB 102 TyrAlaAlaLeuLysAspLeuValLysAlaTyrIleAlaAlaLysGluLysSerAsp 121
QY 373 AACCTTAATGCAACATTAGAACAACTAAATGGCTAAACCTAAATTTAGAATCAGCCATC 432
DB 122 LysSerHisAlaThrLeuGlnGluValAsnAsnAlaLysThrArgLeuGluThrAlaIle 141
QY 433 AACCAAGCTATACGGATAAACAGACTTTTGATAATGAACACCCAAATTTAGTTGAAGA 492
DB 142 LysAspAlaAlaAsnSerLysThrSerPheGlyGluLysAsnProGluLeuIleLysAla 161
QY 493 TACAAAGCCTAAACACACTTTAGAACACACGCTACTACCTTGAAGGTTTGCATCA 552
DB 162 TyrAspAlaLeuLysGlnThrIleThrSerGluGluMetSerLysGlnLeuMetAsp 181
QY 553 ACTGCTTATATCAATTCGCAATTAATTTAGTGGATCTATACATAAAGTAGTAGTTTA 612
DB 182 AlaAsnPheGluThrIleLysAsnHisIleSerAsnLeuTyrLysGlnGlyLysAspIle 201
QY 613 ATAACCTAAACACTAGATCCACTAAATGGGGAGCGCTTTTAGATTCTAATGAG----- 666
DB 202 IleThrAlaThrLeuAspProThrThrGly-----AspGlyProGlnAlaMet 217
QY 667 ---ATTACTACAGCTAAATAAGAAATTAATAATACGTTATCAACTATTAAATGAACAAAG 723
DB 218 ValValAsnGlnThrAsnGluAlaIleValAsnAlaThrSerLysIleGluAspTrpLys 237
QY 724 ACTAATGCTGATGCTATTCTAATAGTTTATTAATAAAGAGTAGTTCAAAAT----- 774
DB 238 ThrAsnAlaThrAsnLeuAlaThrArgPheValLysGlnThrLeuAsnAsnAlaAsnLeu 257
QY 775 ---AATGAACAAGTTTGTAGGGAGCTTTTACAACGCTAATGTTCAACCTTCAACACTAC 831
DB 258 ValAsnGlu-----ThrAsnAsnGlnProGlnProGlySerTyr 270
QY 832 AGTTTGTGCTTTTGTAGTGTGATGTA-----ACACCC 864
DB 271 SerPheValAlaTyrSerValAspLeuAsnThrGlyValSerThrAlaSerAsnThrPro 290
QY 865 GTCATATTATAATATGCAAGAGGACCGTTGG-----AATGGGTGATGAACCT--- 912
DB 291 ---AsnTrpAsnLeuAlaGlnArgLysValTrpValSerGlySerGlyArgThrSer 309
QY 913 -----TCAAGTAGAATTCCTTGCACCAACAGATATCATCAGATGTTTCTTGGATTAT 966
DB 310 ProPheSerSerAspAlaAsnAsnSerProAlaLeuThrAspValSerTrpIleTyr 329
```

```
QY 967 AGTTTAGCTGGAACAAACACACGAGTACCAATTTAGCTTTAGCAACATGATGGTCCATCACT 1026
DB 330 AsnLeuSerGlyAlaAsnSerLysTyrThrLeuThrPheAsnTrpTyrGlyProSerThr 349
QY 1027 GGTATTATTATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTTGGATTA 1086
DB 350 GlyHisLeuTyrPheProTyrLysLeuValLysSerAspAlaGlnAsnValGlyLeu 369
QY 1087 CAATCAAAATTAATAATGGAATGTTCAACAGTTGAGTTGGC-----1131
DB 370 GlnTyrThrLeuAsnAsnLysProAlaGlnArgIleGluPheAlaProAlaGlnSerPro 389
QY 1131 -----1131
DB 390 SerSerGlyGlyThrAlaHisAlaSerAspProGlnSerProArgAlaAlaThrGlu 409
QY 1132 -----ACTTCAACTAGTGCATAATAATACTACTACA 1158
DB 410 ThrAspValSerAspSerAlaGluGlySerGlnAlaGlnThrAspMetSerSerSer 429
QY 1159 GCTAATCCAACTCCAGCAGTTCATCAGATTAAGTTGCTAAATCGTTTATCATCGTTTA 1218
DB 430 MetAsnLysThrProThrValSerAspIleAsnValAlaSerValThrLeuSerAspLeu 449
QY 1219 AGATTGGCCAAACACAAATCGAATTAAGTGTTCACACGGGTGAAGAAATATGAATAA 1278
DB 450 AsnPheGlyAlaAsnThrIleGluPheSerValProMetGlyAsp-----SerMet 466
QY 1279 GTTGGCCCAATGATTGGCAACATTTTCTTACCTCAAAATGAATAATGCTGATGAATC 1338
DB 467 ValAlaProMetIleGlyAsnMetTyrIleThrSerAsnProLeuAsnValAsnGlnIle 486
```

RESULT 13

Q49468

ID Q49468 PRELIMINARY; PRT; 647 AA.
AC Q49468; Q53303;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hemagglutinin homolog precursor.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93162830; PubMed=8432610;
RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmga, a
hemagglutinin of Mycoplasma gallisepticum.";
RL Infect. Immun. 61:903-909(1993).
DR EMBL; M83178; AAA02996.1; --
DR EMBL; S55216; AAB25397.2; --
KW Signal.
FT SIGNAL. 1 25 POTENTIAL.
FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Alignment Scores:

Pred. No.: 2,6e-32 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 2 Gaps: 13

US-09-147-052-1 (1-1371) x Q49468 (1-647)

```
QY 157 GTGGTTCACCGTATCGTATCGCTAGATTGGCTGTCTATCTATTACTAAAAA----- 210
DB 15 IleGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProAsnProThr 34
```



```

Db 183 AspProLysTyrSerAlaIleLeuAspGlnIleAsnGlyValSerSerLysGlyGlu 202
QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCTTAATGAGATT 669
Db 203 LeuValGlnHisThrLeuAspProValSer---GlyIleValProAlaAlaAsnThrIle 221
QY 670 ACTACAGCTAATAAGATATTAATAATACGTTATCA-----ACTATTAAATGAACAAAG 723
Db 222 ThrGluGluIleThrLysIleGluGluValIleSerGluLysThrLeuGlnAspGlnLys 241
QY 724 ACTAATGCTGATGATTCATTAAT-----AGTTTTTAAT-----AAAAAGTATTCAAA 771
Db 242 AsnAsnAlaAspGlnPheAlaAsnTyrGlnSerPheThrLeuAspLysThrLysLeuGlu 261
QY 772 AATAATGAACAAGATTTTGTAGGACCTTTTACAACGCTAATGTTCAACCTTCAACATCAC 831
Db 262 AsnValGluAspAla-----LysLysMetGlyGlnProAlaAsnTyr 275
QY 832 AGTTTGTGCTTTTGTAGTGCTGATGTATACACCCGTC-----LysLysMetGlyGlnProAlaAsnTyr 275
Db 276 SerPheValGlyTyrSerValAspValThrGlyThrSerGlyGlnGluIleThrIlePro 295
QY 868 AATTATAATATGCAAGAACGCTTTTGGATGGTATGATGAACCTTCAAGTAGAATCTT 927
Db 296 AsnTrpAsnPheAlaGlnArgAlaIlePheThrSerGlyAsnGlnProThrLysValThr 315
QY 928 CCAACACGCAATAGT-----ATCAGAGATGTTTCTTGGATT 963
Db 316 AlaThrThrThrGlyGluAspGlnSerThrAlaLysProLeuSerAspValSerTrpIle 335
QY 964 TATAGTTAGCTGGAACAAACACAGATACCAATTTAGTTTATGCAACCTATGTTCCATCA 1023
Db 336 TyrSerLeuAlaGlyThrGlyAlaLysTyrThrLeuGluPheThrTyrTyrGlyProSer 355
QY 1024 ACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTTAATACGTTGGA 1083
Db 356 ThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAsnAsp-----AspValGly 373
QY 1084 TTACAATCAATTAAT---AATGGAATGTTCAACAAAGTTGAGTTGCCACTTCAACT 1140
Db 374 LeuGlnTyrLysLeuAsnSerAsnGluThrLeuThrProIleIlePheGlyGluGlyThr 393
QY 1141 AGTCAAAATAACTACAGCTAATCCAACTCCAGCTGATGAGATTAAGTTAAGTCTAAA 1200
Db 394 -----ThrThrAsnGlyProAlaIleThrValGluAsnIleAsnValAlaLys 409
QY 1201 ATCGTTTATCAGCTTTAAGATTGGCCAAAACACAAATCGAATTAAGTTGTTCCACGCGT 1260
Db 410 ValArgLeuThrGlyLeuAlaPheGlyLysAsnThrIleGluPheSerValPro----- 427
QY 1261 GAAGGAATATCAATAAGTTGCGCAATGATGTCACACATTTATCTTATGCTCAATGAA 1320
Db 428 -----MetSerLysValAlaProMetIleGlyAsnMetTyrIleThrSerSerAsp 444
QY 1321 AATAATGCTGATAAG 1335
Db 445 ThrGluThrAsnLys 449

RESULT 15
O05122 PRELIMINARY; PRT; 703 AA.
AC O05122;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Haenagglutinin.
GN PMGAl-9
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S6;

```

```

RX MEDLINE-95010739; PubMed-7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithead K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U00714; AAB50154.1; -
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CBC5 CRC64;

Alignment Scores:
Pred. No.: 1,76e-31 Length: 703
Score: 700.50 Matches: 170
Percent Similarity: 50.61% Conservative: 77
Best Local Similarity: 34.84% Mismatches: 126
Query Match: 30.12% Indels: 115
DB: 2 Gaps: 12

US-09-147-052-1 (1-1371) x 005122 (1-703)
QY 157 GTGGTTCAACCGTGATCCGCTAGAAATCGCTGTATGCTTACTATAAAAGATGCA 216
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrSerAlaThrIleProThrLeu 34
QY 217 AACCCA----- 222
Db 35 AsnProThrProAsnProGluProLysProAspProMetProAsnProProSerGlyGly 54
QY 223 ---AATAATGCCAAACC-----CAATTAGAACGCG 252
Db 55 MetAsnGlyGlyAsnThrAsnProGlyMetAspThrAlaAlaGlnIleLeuAlaSerAla 74
QY 253 CGAATGGATTAAACAGATCTAATCAATGCTAAGCGATGACATAGCTTCACTCAAGAC 312
Db 75 LysAlaLeuThrThrLeuThrAsnArgGluSerGluLysValGlyLeuTyrValAsp 94
QY 313 TATGCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTAAACAT 372
Db 95 TyrAlaLysIleLysAlaAspLeuThrSerAlaTyrThrValAlaLysThrThrSerAsp 114
QY 373 AACCTTAATGCAACATTAGAACACAACTAAATAATGGCTAAACATAATTTAGATCAAGC 432
Db 115 SerSerThrSerThrLeuValGlnValLysThrAlaThrSerLysLeuGlnThrAlaIle 134
QY 433 AACCAAGCTAATACGGATAAAACGACTTTTGTATATGAAACACCCCAAAATTTAGTTGA 492
Db 135 AspLysAlaAlaSerAspLysGlnLysPheGluGlnAspHisLysAspLeuLeuMetPro 154
QY 493 TACAAAGCACTAAACACCTTTTAGAACAAAGCT---GCTACTAACCTTGAAGTTGTGCA 549
Db 155 TyrSerGluLeuLysThrThrLeuSerGlnLysAsnAlaThrValLeu-----LeuAsn 172
QY 550 TCACTGCTTATATCAAAATTCGCAATTAATTTAGTGGATCTATACATAAAGCTAGTAGT 609
Db 173 GlnProLysTyrSerAlaIleLeuAsnLysIleAsnSerIleTyrAlaGlnGlyGlu 192
QY 610 TTAATACTAAACACTAGATCCCTAAATGGGGAACGCTTTTAGATTCTTAATGAGATT 669
Db 193 ValValIleArgThrLeuAspProValSer---GlyAlaIleProThrAlaAlaSerIle 211
QY 670 ACTACAGCTAATAAGATATTAATAATACGTTATCA-----ACTATTAAATGAACAAAG 723
Db 212 ThrLysValAsnAspGluIleAsnLysAlaIleSerGluAsnGlnLysProLysLys 231
QY 724 ACTAATGCTGATGATTCATTAATAGTTTATTAATAAAAGTCAATTAATAAATGAACAA 783
Db 232 AspAsnAlaAspAlaPheAlaAsnTyrGlnPhePheLysLeu-----AspLysThr 248
QY 784 AGTTTGTAGGACTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTTGTGCT 843

```

```
Db 249 LysIleMetGlyMetSerThrAsnMetMetLysGlnProGlnAsnTyrSerPheValGly 268
QY 844 TTTAGTCTGATGTAACACCGTC-----AATTATAATAT 879
Db 269 TyrSerValGlyValThrGlyMetGlnSerGlyGlnThrThrIleProAsnTyrAsnPhe 288
QY 880 GCAAGAAAGGACCGTTTGGATGGTGATGAACCTTCAAGTAGAATTCTTGCACACACGAAT 939
Db 289 AlaglnArgIleValTrpSerSerGlyAlaProArgAlaProLeuAlaSerGlnThrGlu 308
QY 940 AGT----- 942
Db 309 ThrProGlnAlaGluThrProMetSerAlaProGlnGlyValGluProAlaGlnGln 328
QY 942 ----- 942
Db 329 GlnGlyAspSerSerProLysGlnAlaSerGluThrGlnGluValSerProThrProAla 348
QY 942 ----- 942
Db 349 AlaGluValGlnAlaGlnGlnAlaAspThrGluGlnProAlaThrSerGlnGlyThrPro 368
QY 943 ATCACAGATGTTCTTGGATTATTAGTTAGCTGGAACAACACGAAAGTACCAATTTAGT 1002
Db 369 LeuThrAspValSerTrpIleTyrSerLeuSerGlyThrAspValLysTyrThrPheThr 388
QY 1003 TTTAGCAACTATGCTCCACTGCTGTTATTTATATTTCCCTTATTAAGTTGGTTAAAGCA 1062
Db 389 PheAsnTyrPheGlyProSerMetAlaTyrLeuTyrPheProTyrLysLeuValLysSer 408
QY 1063 GCTGATGCTAATAACGTTGATTACAAATCAAAATTAATAATGAAATGTTCAACAAGTT 1122
Db 409 AspAsp-----SerValGlyLeuGlnTyrLysLeuAsnAsnAsnProValAlaLeu 426
QY 1123 GACTTTCCCACTTCAACTAGTCGAAATATATCTACAGCTAATCCAACTCCAGCAGTTGAT 1182
Db 427 AsnPheGlySerGluThrAsnAlaAsn-----GlyProAlaAlaSerValAsp 442
QY 1183 GAGATTAAAGTTGCTAAATCGTTTATCAGGTTTAAAGATTTGGCCAAACACATCGAA 1242
Db 443 AsnIleAsnValAlaLysValAsnLeuAlaAsnLeuAsnPheGlyGluAsnThrIleGlu 462
QY 1243 TTAAGTGTCCCAACGGGTGAAGGAAATATCAATAAGTTCCGCCAATGATTGGCAACATT 1302
Db 463 PheSerValPro-----MetAsnLysValAlaProMetIleGlyAsnMet 477
QY 1303 TATCTTAGCTCAAAATCAAAATAT 1326
Db 478 TyrIleThrSerAspValAlaAsn 485
```

Search completed: October 8, 2003, 18:10:56
Job time : 141.423 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:33:17 ; Search time 22.0964 Seconds
(without alignments)
5835.661 Million cell updates/sec

Title: US-09-147-052-1

Perfect score: 2326
Sequence: 1 atgcaactatttagcgaggaa.....gaccocgtacattttataa 1371

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q/cgn2.1/USPTO_spool/US09147052/runat_08102003_154340_29835/app_query.fasta_1.5980
-DB=SwissProt_41 -OPMT=fastan -SUFFIX=esp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MAPRI=blosum62 -TRANS=human40 cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09147052@cgn_1.140.0@runat_08102003_154340_29835 -NCPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.1*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	314	13.5	865	1 VGLB_HSVMD	P18338 maret's dis
2	154.5	6.6	682	1 NISP_LACLA	Q07596 lactococcus
3	152	6.5	2660	1 YEF1_ECO57	Q8x8v7 escherichia
4	146.5	6.3	857	1 AR56_CANAL	P78586 c arg5.6 pr
5	146.5	6.3	1085	1 CUV7_SCHPO	P24339 schizosacch
6	146	6.3	719	1 YM41_YEAST	Q03213 saccharomyc
7	143	6.1	3178	1 YS89_CAEEL	Q09624 caenorhabdi
8	141	6.1	1314	1 SW11_YEAST	P09547 saccharomyc
9	140	6.0	1381	1 YB7_YEAST	P34216 saccharomyc
10	138	5.9	444	1 PSY1_YEAST	Q12355 saccharomyc
11	138	5.9	995	1 Y109_YEAST	P40442 saccharomyc
12	138	5.9	1010	1 YK11_CAEEL	P34278 caenorhabdi
13	138	5.9	1260	1 ALS1_CANAL	P46590 candida alb
14	137	5.9	1140	1 YM96_YEAST	Q04893 saccharomyc
15	136.5	5.9	956	1 YEF3_YEAST	P32618 saccharomyc
16	135	5.8	1271	1 Y338_MYCGE	P47580 mycoplasma
17	133.5	5.7	903	1 VGLB_HSV1F	P06436 herpes simp
18	133.5	5.7	904	1 VGLB_HSV11	P10211 herpes simp

19	133.5	5.7	904	1 VGLB_HSV1P	P08565 herpes simp
20	133.5	5.7	1018	1 FNBA_STAAU	P14738 staphylococ
21	132.5	5.7	457	1 MESE_LEUME	Q10419 leuconostoc
22	132.5	5.7	1672	1 PMPB_CHLMU	Q9pjy2 chlamydia m
23	132	5.7	1902	1 P3P_LACLC	P15392 lactococcus
24	131.5	5.7	1164	1 BAG_STRAG	P27951 streptococ
25	131	5.6	837	1 RO01_YEAST	Q02805 saccharomyc
26	131	5.6	1433	1 CAT8_YEAST	P39113 saccharomyc
27	130.5	5.6	704	1 MSN2_YEAST	P33748 saccharomyc
28	130.5	5.6	800	1 INLA_LFSMO	P25146 listeria mo
29	130.5	5.6	1159	1 N124_SCHPO	Q09904 schizosacch
30	130.5	5.6	1161	1 DAN4_YEAST	P47179 saccharomyc
31	130.5	5.6	1902	1 P2P_LACLC	P15293 lactococcus
32	130	5.6	425	1 YB10_YEAST	P38272 saccharomyc
33	130	5.6	750	1 YK57_YEAST	P34231 saccharomyc
34	129	5.5	1251	1 RBP2_PLAVB	Q00799 plasmodium
35	129	5.5	1790	1 US01_YEAST	P25386 saccharomyc
36	128.5	5.5	1858	1 P3K2_DICDI	P54674 dictyostell
37	128.5	5.5	2334	1 WAPA_BACSU	Q00783 bacillus su
38	128.5	5.5	3712	1 LMA_DROME	Q00174 drosophila
39	128	5.5	879	1 MYSP_DROME	P35415 drosophila
40	128	5.5	1117	1 YN96_YEAST	P53753 saccharomyc
41	127.5	5.5	544	1 FLGK_BUCAP	Q8k9k0 buchnera ap
42	127	5.5	762	1 SLAP_ACEKI	P22358 acetogenium
43	127	5.5	979	1 P115_MYCHR	P41308 mycoplasma
44	127	5.5	1142	1 GINA_YEAST	Q12263 saccharomyc
45	127	5.5	1643	1 OMPB_RICPR	Q53020 r outer mem

ALIGNMENTS

RESULT 1

ID	VGLB_HSVMD	STANDARD;	PRT;	865 AA.
AC	P18538;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	GB.			
OS	Marek's disease herpesvirus (strain RB-1B) (MDHV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_Taxid=33707;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89293086; PubMed=2544666;			
RA	Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;			
RT	'Nucleotide sequence and characterization of the Marek's disease			
RT	virus homologue of glycoprotein B of herpes simplex virus.';			
RL	J. Gen. Virol. 70:1789-1804(1989).			
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D13713; BAA02866.1; -			
DR	InterPro; IPR000234; Glycoprot.B.			
DR	Pfam; PF00606; Glycoprotein_B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			
KW	Signal; Glycoprotein; Transmembrane.			
FT	SIGNAL 1 21 POTENTIAL.			
FT	CHAIN 22 865 GLYCOPROTEIN B.			
FT	DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 683 700 POTENTIAL.			
FT	TRANSMEM 709 729 POTENTIAL.			
FT	TRANSMEM 732 752 POTENTIAL.			
FT	DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).			

FT CARBOHYD 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 98091 MW; B30E93C1AC65C6C3 CRC64;

Alignment Scores:
 Pred. No.: 2.02e-11 Length: 865
 Score: 314.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.50% Indels: 0
 DB: 1 Gaps: 0

US-09-147-052-1 (1-1371) x VGLB_HSVMD (1-865)

QY 1 ATGCACATTTTAGCGGAATTCATATTTTCCTTATAGTATATCTATATGTCAGAAC 60
 Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCTCCAGTACCCAAATGTGACATCAAGAGAAGTGTTCGAGCGTCCAGTGTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTACGTTTATCTTCTTGTCGCCACAGTGGTTCACCGTGATCCGTCTA 180
 Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAA 183
 Db 61 Glu 61

RESULT 2

NISP_LACLA STANDARD; PRT; 682 AA.
 AC Q07596;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nisin leader peptide processing serine protease nisp precursor
 DE (EC 3.4.21.-).
 GN NISP.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIZO R5;
 RA MEDLINE=93239683; PubMed=8478324;
 RA van der Meer J.R., Polman J., Beerthuyzen M.M., Sierzen R.J.,
 RA Kuipers O.P., de Vos W.M.;
 RT "Characterization of the Lactococcus lactis nisp A operon genes
 nisp, encoding a subtilisin-like serine protease involved in
 precursor processing, and nispR, encoding a regulatory protein
 involved in nisin biosynthesis.";
 RL J. Bacteriol. 175:2578-2588(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6F3;
 RA MEDLINE=94213459; PubMed=8161176;
 RA Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,
 RA Hammelmann M., Entian K.-D.;
 RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
 6F3.";
 RL Appl. Environ. Microbiol. 60:814-825(1994).
 RN [3]
 RP 3D-STRUCTURE MODELING.
 RA MEDLINE=95357326; PubMed=7630881;
 RA Slezien R.J., Rollena H.S., Kuipers O.P., de Vos W.M.;
 RT "Homology modelling of the Lactococcus lactis leader peptidase Nisp
 and its interaction with the precursor of the lantibiotic nisin.";

Protein Eng. 8:117-125(1995).
 CC -!- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
 CC -!- PATHWAY: Nisin biosynthesis; last step.
 CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 an amide bond (Potential).
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

EMBL; L11061; AAA25200.1; -;
 EMBL; X76884; CAA54210.1; -;
 DR PIR; S44131; S44131.
 DR HSSP; P29600; 1GCI.
 DR MEROPS; S08.059; -;
 DR InterPro; IPR006192; LPXTG.
 DR InterPro; IPR002029; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
 DR PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
 KW Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
 KW Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 195 POTENTIAL.
 FT CHAIN 196 655 NISIN LEADER PEPTIDE PROCESSING SERINE
 FT PROPEP 656 682 PROTEASE NISP.
 FT ACT_SITE 259 259 REMOVED BY SORTASE (POTENTIAL).
 FT ACT_SITE 306 306 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 512 512 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT SITE 652 656 LPXTG SORTING SIGNAL (POTENTIAL).
 FT MOD_RES 655 655 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 FT CONFLICT 500 500 A -> T (IN REF. 2).
 SQ SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Alignment Scores:

Pred. No.: 0.036 Length: 682
 Score: 154.50 Matches: 75
 Percent Similarity: 38.10% Conservative: 61
 Best Local Similarity: 21.01% Mismatches: 130
 Query Match: 6.64% Indels: 91
 DB: 1 Gaps: 13

US-09-147-052-1 (1-1371) x NISP_LACLA (1-682)

QY 358 GAAACAGTAAACAATAACCTTATGCAACATTAGAACAACTAAATAATGCTAAATAAT 417
 Db 38 GluLeuIleAsnHisAsnSerAsnAlaIleLeuSerThrGluGlySerThrThrAsp 57
 QY 418 TTAGATATCATCAACCAAGCTATACGGATATACGGAATTAACGACCTTTTGATATGACACCCA 477
 Db 58 SerIleAsnLeuGlyAlaGlnSerProAlaValLysSerThrThrArgThrGlu----- 75
 QY 478 AATTAGTTGAAGCATACAAAGCACTTAAAGAACCTTTAGAACAACTGCTACTAACCTT 537
 Db 76 -----LeuAspValThrGlyAlaAlaLysThrLeuLeuGlnThrSerAlaValGlnLys 93
 QY 538 GAAGGTTTTCATCAACTGCTTATATCAATTCGAATAATTAATAGTGATCATACAT 597
 Db 94 GluMetLysValSerLeuGlnGluThrGlnValSerSerGlu-----PheSer 109
 QY 598 AAAGCTAGTAGTTTAAATAACTAAACACTAGATCCACTAAATGGGGGACCGCTTTAGAT 657
 Db 110 LysArgAspSerValThrAsnLysGluAlaValProValSerLysAspGluLeuGlu 129

```

QY 658 TCTAATGAGATTACTACAGCTAATAAGAAATATTATAATACGTATCACTAATTAATGAA 717
Db 130 GlnSerGluValValSerThrSerIleGlnLysAsn---LysIleLeuAspAsn 148
QY 718 CAAGAAGCTAATGCTGATGCTATCTAATAGTTTATTAATAAAGAGTATTCAAAATAT 777
Db 149 LysLysArgAlaAsnPheValThrSerProLeuIleLysGluLysProSerAsn 168
QY 778 GAACAAAGCTTTGTAGGAGCTTTTACAAACGCTTAATGTTCAACCTTCAACTACAGTTT 837
Db 169 SerLysAspAlaSerGlyValIleAspAsnSer----- 179
QY 838 GTTGCTTTTAGCTGATGCTACACCCGCTCAATATAATATGCAAGAGGACCGTT--- 894
Db 180 -----AlaSerProLeuSerTyrArgLysAlaLysGluValValSer 193
QY 895 -----TGGAAATGCT 903
Db 194 LeuArgGlnProLeuLysAsnGlnLysValGluAlaGlnProLeuLeuIleSerAsnSer 213
QY 904 GATGAACCTTCAAGPAGAAATCTTGCACACGAAATAGTATCACAGATGTTCTTGGATT 963
Db 214 SerGluLysLysAlaSerValTyrThrAsnSerHisAspPheTyrAspTyrGlnTrp--- 232
QY 964 TATAGTTTACGCTGGCAACACAGCAAGTACCAATTTAGTTTACCACTATGGTCCATCA 1023
Db 233 -----AspMetLysTyr-----ValThrAsnAsnGlyGluSer 243
QY 1024 ACTGCTATTATATTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACCTTGA 1083
Db 244 TyrAla---LeuTyrGlnProSerLysLysIle-----SerValGly 256
QY 1084 TTACAATACAATAAATGAATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGT 1143
Db 257 Ile-----IleAspSerGlyIleMetGluGluHisProAspLeuSerAsnSerLeu 273
QY 1144 GCAAT-----AATACT 1155
Db 274 GlyAsnTyrPheLysAsnLeuValProLysGlyGlyPheAspAsnGluProAspGlu 293
QY 1156 ACAGCTAATCAACATCCAGCAGCTGATGAGATTAAGTTGCTAAATCGTTTATCAGGT 1215
Db 294 ThrGlyAsnProSerAspIleValAsp----- 302
QY 1216 TTAAGATTGGCCAAACACATCAATCAATTAAGTTTCCACGCGGTGAGGAATATGAAT 1275
Db 303 ---LysMetGlyHisGlyThrGluValAlaGlyGlnIleThrAlaAsnGlyAsnIleLeu 321
QY 1276 AAGTTGCGCCCAATATTGGC---AACATTATCTTAGCTCAAAATGAAAT 1323
Db 322 GlyValAlaProGlyIleThrValAlaSerIleTyrArgValPheGlyGluAsn 338

RESULT 3
YEEJ_ECO57
ID Q8X8V7; Q8X2B9; Q8X2C0; PRT; 2660 AA.
AC Q8X8V7; Q8X2B9; Q8X2C0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yeeJ.
GN Z3135 OR ECS2775/ECS2776.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoumis K.,

```

```

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC - SIMILARITY: Contains 16 Big-1 domains.
CC - SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.
CC - CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 1315.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005423; AAG57041.1;
DR EMBL; AP002559; BAB36198.1; ALT_FRAME.
DR EMBL; AP002559; BAB36199.1; ALT_FRAME.
DR InterPro; IPR003344; Big_1.
DR InterPro; IPR003535; Intimin.
DR Pfam; PF02369; Big_1; 16.
DR PRINTS; PR01369; INTIMIN.
DR SMART; SM00634; BID_1; 16.
DR SMART; SM00089; PKD; 8.
DR KW Hypothetical protein; Repeat; Complete proteome.
FT DOMAIN 738 834 BIG-1 1.
FT DOMAIN 840 929 BIG-1 2.
FT DOMAIN 931 1033 BIG-1 3.
FT DOMAIN 1042 1132 BIG-1 4.
FT DOMAIN 1134 1236 BIG-1 5.
FT DOMAIN 1245 1335 BIG-1 6.
FT DOMAIN 1337 1439 BIG-1 7.
FT DOMAIN 1448 1539 BIG-1 8.
FT DOMAIN 1548 1652 BIG-1 9.
FT DOMAIN 1653 1750 BIG-1 10.
FT DOMAIN 1751 1855 BIG-1 11.
FT DOMAIN 1856 1957 BIG-1 12.
FT DOMAIN 1963 2056 BIG-1 13.
FT DOMAIN 2065 2156 BIG-1 14.
FT DOMAIN 2157 2252 BIG-1 15.
FT DOMAIN 2254 2355 BIG-1 16.
SQ SEQUENCE 2660 AA; 280062 MW; 01EB92A08F5C09D2 CRC64;

Alignment Scores:
Pred. No.: 0.0447 Length: 2660
Score: 152.00 Matches: 108
Percent Similarity: 37.87% Conservative: 70
Best Local Similarity: 22.98% Mismatches: 188
Query Match: 6.53% Indels: 104
DB: 1 Gaps: 20

US-09-147-052-1 (1-1371) x YEEJ_ECO57 (1-2660)
QY 34 CUTATAGTATTCTATATGTTACGAGTACATCGAGTACCCAAATGTGATCAAGA 93
Db 1131 IleValAsnIleAlaProAspAlaSerAsnAlaGlnValThrLeuAsnIleProAlaGln 1150
QY 94 GAAGTTGTTTCG-----AGCGTCCAGTTGTCT-----GAGGAAGAGTCT 132
Db ::::|||||:

```

Db 1151 GlnValValThrAsnAsnSerAspSerValGlnLeuThrAlaThrValLysAspProSer 1170
 QY 133 ACGTTTATCTTGTCCCCACAGTGGGTTCAACCGTGATCCGCTAGAAATCGCGCTGT 192
 Db 1171 AsnHis-----ProValAlaGlyIleThrVal----- 1179
 QY 193 ATGCTATTACTAAAGAGTCAAAACCCCAAAATAATGGCCCAACCCCAATAGAGCAGCG 252
 Db 1180 -----AsnPhetThrMetProGlnAspValAlaAla 1189
 QY 253 CGAATGGAGTTAACAGATCTTAATCAATGCTAAAGCGATGATGATCTACCTACACAGC 312
 Db 1190 AsnPhetThrLeuGluAsn-----AsnGlyIleAlaIleThrGlnAlaAsnGlyGluAla 1207
 QY 313 TATGCCAAGATTCAAGCTAGTATTATCATCTGCTTTATAGTGAAGCTGAACAGTAAACAAT 372
 Db 1208 HisValThrLeuLysGlyLysAlaGlyThrHisThrValThrAlaThrLeuGlyAsn 1227
 QY 373 AACCTTAATGCAACATTAGAACAACTAAATAAGCTTAATAACTAATTTAGAACTCAGCCATC 432
 Db 1228 Asn---AsnAlaSerAspAlaGlnProValThrPheValAlaAspLysAspSerAlaVal 1246
 QY 433 AACCAAGCTAATACG-----GATAAAAGCACT--- 459
 Db 1247 ValValLeuGlnThrSerLysAlaGluIleIleGlyAsnGlyValAlaGluThrThrLeu 1266
 QY 460 -----TTTGATATGAACACCCCAATTTAGTTGAAGCATACAAA 498
 Db 1267 ThrAlaThrValLysAspProPheAspAsn-----AlaValLys 1279
 QY 499 GCCTAAACCACTTTAGAACACAGCTGCTACTAACTTGAAGTTTGTTCATCACTGCT 558
 Db 1280 AspLeuGlnValThrPhe-----SerThrAsn-----ProAlaAspThrGln 1293
 QY 559 TATAATCAAAATCGCAATTAATTTAGTGGATCTATACAATAAAGCTAGTATTAAACT 618
 Db 1294 LeuSerGlnSerLysSerAsnThrAsnAsp-----SerGlyValAlaGlu 1308
 QY 619 AAACACTAGATCCACTAAATGGGGACGCTTTTA-----CATCTAATGAGATTACT 672
 Db 1309 ValThrPheLys-----GlyThrValLeuGlyValHisThrAlaGluAlaThr 1324
 QY 673 ACAGCTAATAAGAAATTAATAATAGCTTATCAACTATTAAATGAACAAAAGACTAATGCT 732
 Db 1325 LeuProAsnGlyAsnAsnAspThrLysIleValAsnIleAlaProAspAlaSerAsnAla 1344
 QY 733 GATGCATTATCTAATAGTTTATTAATAAAGTGATTCAAAATTAATGAACAAAGTTTGA 792
 Db 1345 GlnValThrLeuAsnIleProAlaGlnGlnValValThrAsnAsnSerAspSerValGln 1364
 QY 793 GGCACCTTTTACAACGCTAATGTTCAACCTTCAACTACAGTTTGTGCTTTTAGTGCT 852
 Db 1365 LeuThrAlaThrValLysAsp-----ProSerAsnHisProValAlaGlyIleThrVal 1382
 QY 853 GATGTAAACCCGCTCAATTATTAATAATATGCAAGAGGACCGTTTGGAAATGCTGATGAACCT 912
 Db 1383 AsnPhetThrMetProGlnAspValAlaAlaAsnPhetThrLeu-----GluAsn 1398
 QY 913 TCAAGTAGAATTCCTGCAACACGCAATAGTATGATCAGAGATTTCTTGGATTATAGTTTA 972
 Db 1399 AsnGlyIleAlaIleThrGlnAlaAsnGlyGluAlaHisValThrLeuLysGlyLysLys 1418
 QY 973 GCTGGAACAACACAGTACCAATTTAGTTTATGCAACTATGCTCCACTCACTGGTTAT 1032
 Db 1419 AlaGlyThrHisThr---ValThrAlaThrLeuSerAsnAsnAsnThrSerAspSerGln 1437
 QY 1033 TTATATTCCTTATTAAGTTGGTTAAGCAGCTGATGCTAATAACGCTGGATTACATAATC 1092
 Db 1438 -----ProValThrPheValAlaAspLysThrSerAlaLeuValValLeuGlnIle 1454
 QY 1093 AAATTAATAATGGAATGTTCAACAAGTTGAGTTGGCCACT----- 1134
 Db 1455 SerLysAsnGluIleThrGlyAsnGlyValAspSerAlaThrLeuThrAlaThrValLys 1474

QY 1135 -----TCAACTAGTGCATAAAT 1152
 Db 1475 AspGlnPheAspAsnGluValAsnAsnLeuProValThrPheSerThrAlaSerSerGly 1494
 QY 1153 ACTACAGCTAATCAACTCCAGCA---GTTGATGAGATTAAAGTTGCTAAATCGTTTAA 1209
 Db 1495 LeuThrProGlyGluSerAsnThrAsnGluSerGlyIleAlaGlnAlaThrLeu 1514
 QY 1210 TCAGTTTAAAGATTGGCCAAACACAAATCAATTAAGTTGTTCCACGGGTGAAGGAAAT 1269
 Db 1515 AlaGlyValAlaPheGlyGluGlnThrValThrAlaSerLeuAlaAsnAsnGlyAlaSer 1534
 QY 1270 ATGAATAAAGTTGGCCCAATGATGTCGCAAC 1299
 Db 1535 AspAsnLysThrValHisPheIleGlyAsp 1544
 RESULT 4
 AR56_CANAL STANDARD; PRT; 857 AA.
 AC P78586;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ARG5,6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-L-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-phosphotransferase)].
 GN ARG5,6.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OC NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 64385 / 1001;
 RX MEDLINE=97195775; PubMed=9043106;
 RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
 RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of Candida albicans";
 RT Microbiology 143:297-302(1997).
 RL
 CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+) + phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
 CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-glutamate 5-phosphate.
 CC -1- PATHWAY: Arginine biosynthesis; second step.
 CC -1- PATHWAY: Arginine biosynthesis; third step.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ACETYLGUTAMATE KINASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X98880; CAA67383.1; -.
 CC InterPro; IPR001048; Aa_kinase.
 CC InterPro; IPR004662; AcgLuKinase.
 CC InterPro; IPR000706; AGPR_act_site.
 CC InterPro; IPR006855; DUF619.
 CC InterPro; IPR000534; Semialdh_dh.
 CC Pfam; PF00696; akinase; 1.
 CC Pfam; PF04768; DUF619; 1.
 CC Pfam; PF01118; Semialdehyde_dhc; 1.
 CC Pfam; PF02774; Semialdehyde_dhc; 1.
 CC ProDom; PD003765; AGPR_act_site; 1.

TIGREMS: TIGR00761; argB; 1.
 DR PROSITE: PS01224; ARGC; 1.
 KW Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
 KW Mitochondrion; Multifunctional enzyme; Transit peptide.
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? ? ACETYLGUTAMATE KINASE.
 FT CHAIN ? 857 N-ACETYLGAMMA-GLUTAMYL-PHOSPHATE
 FT REDUCTASE.
 FT ACT_SITE 669 669 BY SIMILARITY.
 SO SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C4C71 CRC64;

Alignment Scores:
 Pred. No.: 0.103 Length: 857
 Score: 146.50 Matches: 122
 Percent Similarity: 35.61% Conservative: 71
 Best Local Similarity: 22.51% Mismatches: 186
 Query Match: 6.30% Indels: 163
 DB: 1 Gaps: 32

US-09-147-052-1 (1-1371) x AR56_CANAL (1-857)

Qy	19	AATTGCATATTTTCCCTTATA-----GTTATCTATATGGTACGAAC	60
Db	104	SerCysLeuAlaPheLeuTyrHisValGlyLeuTyrProIleValLeuHisGlyThr----	122
Qy	61	TCATCTCCGACTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGCT	120
Db	123	---GlyProGlnIleAsnGluLeuGluAsnGluGlyValGluProGluTyrIleAsp	141
Qy	121	GAGGAAGAGTCTAGCTTTATCTTTGTCCTCCACCGAGTGGTTCACCGTGATCCGT---	177
Db	142	GlyIleArgIleThr-----AsnProLysThrMetGluValValArgLys	156
Qy	178	-----CTAGAATTCGGCTGTATCTATTACTAAAAAGAT-----GCA	216
Db	157	CysPheLeuGluGlnAsnLeuArgLeuValThrAlaLeuGluLysIleGlyValHisAla	176
Qy	217	AACCAATATATGCGCAACCCATTAGAACCGCGGAATGGAGTTAACAGATCTAATC	276
Db	177	ArgPro-----IleThrAlaGlyValPheGluAlaGluTyrLeuAsp	190
Qy	277	AATGCTAAAGGATGACATTAGCTTCACTACAGACTATGCCAAG-----ATTGAAGCT	330
Db	191	LysAspLysTyrGlnLeuValGlyLysIleThrSerValAsnLysSerProValGluAla	210
Qy	331	AGTTTATCATCTGCTTAT-----AGTGAAGCTGAACAGTTTAACAAT---	372
Db	211	AlaIleAsnSerGlyTyrLeuProIleLeuThrSerLeuAlaGluThrSerSerGlyGln	230
Qy	373	-----AACCTTATGCACATTA-----GACCACTA	399
Db	231	LeuLeuAsnValAsnAlaAspValAlaAlaGlyGluLeuAlaArgGluPheGluProLeu	250
Qy	400	AAATGGCTAAACTAATTTAGATACGCCATCAACCAAGCTAATACGGATAAA-----	453
Db	251	LysIleValTyrLeuAsnGluLysGlyGlyIleIleAsnGlyAsnThrGlyGluLysVal	270
Qy	454	-----ACGACTTTTGATAATGAACCCCAATTTAGTT-----GAAGCATACAAAGCACTA	504
Db	271	SerAlaIleAsnLeuAspGluGluTyrGluAspLeuLeuLysGluSerTipValLysTyr	290
Qy	505	AAACCACTTTAGACACACGT-----GCTACTACCTTGAAGGTTTGTC	549
Db	291	GlyThrLysLeuLysIleLysGluIleHisAspLeuLeuGlnHisLeuProArgSerSer	310
Qy	550	TCAACTGCTTATATCAATTCGCAATATTTAGTGGATCTATACAATAAAGCTAGTAGT	609
Db	311	SerValAla-----IleIleAspValAsnAspLeuGlnLysGlu	323
Qy	610	TTAATAACTAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTTAATCAGATT	669
Db	324	LeuPheThrAspSer-----GlyAlaGlyThrLeu-----Ile	334

Qy	670	ACTACAGCTATAAGAATATTAAT---AATACGTTTATCACTATTATTAATGAACAAAGACT	726
Db	335	ArgArgGlyTyrArgLeuIleAsnArgAsnSerLeuArgAspPheGly-----	350
Qy	727	AATGCTGATGATTCATTAATAGTTTATTAATAAAGTGATTCAAAATAATGAACAAAGT	786
Db	351	AsnProAspLeuLeuArgAsnAlaLeuLeuArg-----	361
Qy	787	TTTGTAGGAGCTTTTACAAACGCTAATGTTCACACTTCAAACTACAGTTTGTGCTTTT	846
Db	362	-----AspProGluIleLysThrGlyLysValSerValAlaSerTyr	375
Qy	847	AGTGCTGATGTAAACACCGTCAATATATAAATATGCAAGAAGCAGCTTTGGAATGGTGAT	906
Db	376	LeuLysPheLeuAspSerValGlnPheLysSer-----TyrGlyAsp	389
Qy	907	GAACCTTCAAGTAGAATTTCTTGCACAAACAG-----	936
Db	390	GluPro---LeuGluValLeuAlaIleValGluGlnAsnAspLysIleProLysLeu	408
Qy	937	-----AATAGTATCATCAGAT-----	951
Db	409	AspGluPheLeuSerSerLysThrGlyTrpLeuAsnAsnValThrAspAsnIlePheAsn	428
Qy	952	-----GTTCTTGGATTTATAGTTTAGCTGGGAACAACACG	987
Db	429	AlaIleLysLysAspTyrSerGlnLeuCysTrpValValAsnGluAsnAspAlaAsnLeu	448
Qy	988	AAGTACCAATTTAGTTTTCAGCACTATGCT---CCATCAACTGGTTATTATATATTCCT	1044
Db	449	ProTrpTyrPheSerLysSerAspGlySerPheAlaLysAsnGlyGlnIleLeuPheTrp	468
Qy	1045	TAT-----AAGTTGGTTAAAGCAGCTGATGCTCAATAACGTT	1080
Db	469	TyrGlyLeuAsnIleAspGluAlaSerLysLeuIleLysGluPheAspSerSerSerIle	488
Qy	1081	GGA-----TTACATACAAATTAATAATGGA-----AATGTTCAACAAGTTGAG	1125
Db	489	GlySerSerLeuSerSerSerLysGluSerGlyValPheThrSerAlaGlnGlnLysArg	508
Qy	1126	---TTTGGCACTTCAACTAGTGAATAATACTACAGCTAATCCAACTCCAGACGTTGAT	1182
Db	509	GlyPheHisHisSerThrValArgArgAsnThr-----AsnProAsnProProLeuSer	526
Qy	1183	GAGATTAAAGTTGCT-----AAAATCGTTTATCAGGTTTAAAGA-----TTTGGC	1227
Db	527	GluGlyLysGlnThrGluArgLysLysValAlaLeuIleGlyAlaArgGlyTyrThrGly	546
Qy	1228	CAAAACACAAATCGAATTAAGTTTCCCAACGGGTGAGGAATATGAATAAAGTTCCGCCA	1287
Db	547	GlnAsnLeuIleLysLeuIle-----AspAsnHisPro	557
Qy	1288	ATGATTGGCAACATTTATCTTAGCTCAAAATGAAATAATGCTGATAAGATCCCCGGTAC	1347
Db	558	TyrLeuAspIleSerTyrValSerSerArgGluLeuGluGlyGlnLysLeuGlnGlyTyr	577
Qy	1348	CGTCGA	1353
Db	578	AsnLys	579

RESULT 5
 CUT7_SCHPO
 ID CUT7_SCHPO STANDARD; PRT; 1085 AA.
 AC P24339;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein cut7.
 GN CUT7 OR SPAC25G10.07C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.


```

QY 736 GCATTA-----TCTAATAGTTTATTATAAAA-----GTGATT 768
Db 823 TleLeuHisSerHisLeuAsnAspThrAsnSerAsnIleArgLysAlaAsnGluIleMet 842
QY 769 CAATAATGAACAAAGTTTGTAGGAGCTTTTCAACACGCTAATGTTTCAACCTTCAAC 828
Db 843 AsnAsnArgSerGluGluPheLeu-----ArgAsnAla----- 853
QY 829 TACAGTTTGTGCTTTAGCTGCTGANGTAACCCGTCATATATAATATGCAAGAAG 888
Db 854 -----AlaSerGlnAlaGluIleValGlyAlaAsnLysGluArgIleGlnLys 869
QY 889 ACCGTTTGAATGCT-----GATCAACCTTCAAGTACAGTTCCTGCAACACG 936
Db 870 ThrValGluAsnGlySerGlnLeuLeuAspSerLysSerLysAlaIleHisSerAsnSer 889
QY 937 AATAGTATCACAGATGTTTCTGATTTATAGTTAGCT----- 975
Db 890 ArgSerMetTyrAsp-----HisCysLeuAlaLeuAlaGluSerGlnLysGln 905
QY 976 GGACAAACACAGACTACAA-----CAATACAAATTAATAATGGAATGTTCAA 1116
Db 906 GlyValAsnLeuGluValGlnThrLeuAspArgLeuLeuGlnLysValLysGluHisSer 925
QY 1003 TTACCAACTATGTCCTCACTACCTGCTTATTTATATTTCCCTTATAGTTGTTAAAGCA 1062
Db 926 GluAspAsnThrLysGluLysHisGlnGlnLeu-----LeuAspLeuLeuGluSer 942
QY 1063 GCTGATGCTAATACGTTGGATTA-----CAATACAAATTAATAATGGAATGTTCAA 1116
Db 943 LeuValGlyAsnAsnAspAsnLeuIleAspSerLysThrProHisThrGluLeuGln 962
QY 1117 CAAGTTGAG-----TTTGCCACTTCAACTAGTGCATAATTAATCTACA----- 1158
Db 963 LysIleThrAspHisValLeuLysGlyThrThrSerLeuAlaAsnHisThrAsnGluLeu 982
QY 1159 -----GCTAATCAACTCCAGCTGATGATGATGATTAAGTTT 1194
Db 983 LeuGlyLeuGlyAspGluSerLeuCysAsnLeuGluThrThrIleGluAspThrSerLeu 1002
QY 1195 GCTAAATCGTTTATCATAGTTTAAAGTTTGGCAAAACAAATC----- 1239
Db 1003 ValLysLeuGluThrThrGlyAspThrProSerLysArgGluLeuProAlaThrProSer 1022
QY 1240 -----GAATTAAGTGTTCACACGGGTGAAGAAATATGAATAAGTTCCG 1284
Db 1023 TrpThrArgAspSerSerLeuIleLysGluThrThrAsnLeuAsnLeuAspSerAspLys 1042
QY 1285 CCAATGATGCAACATTTATCTAGCTCAATGAAATATATGCTGATAAGATCCCCGG 1344
Db 1043 LysPheValArgGluThrThrTyThrSerSerAsnGlnThrAsnGluProAspVal----- 1060
QY 1345 TACCGTCGACCC 1356
Db 1061 TyrAspLysPro 1064

RESULT 6
Y41_YEAST
ID Y41_YEAST STANDARD; PRT; 719 AA.
AC Q03713.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 79.4 kDa protein in ALD2-DDR48 intergenic region.
GN YMR172W OR YMR010.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;

```

```

RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).
CC -1- SIMILARITY: LOW, TO YEAST MSN1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49808; CAA89905.1; -
DR PIR; S51119; S51119.
DR TRANSFAC; T04601; -.
DR SGD; S0004783; HOT1.
DR GO; GO:0000228; C:nuclear chromosome; IDA.
DR GO; GO:0006972; P:hyperosmotic response; IGI.
KW Hypothetical protein.
SQ SEQUENCE 719 AA; 79415 MW; 4652BE93743D5A54 CRC64;
Alignment Scores:
Pred. No.: 0.112 Length: 719
Score: 146.00 Matches: 101
Percent Similarity: 36.02% Conservative: 69
Best Local Similarity: 21.40% Mismatches: 178
Query Match: 6.28% Indels: 128
DB: 1 Gaps: 20
US-09-147-052-1 (1-1371) x Y41_YEAST (1-719)
QY 55 ACRAACTATCTCCGAGTACCCAAATGTGACATCAAGAGAGTTGTTTCGAGCTCCAG 114
Db 150 ThrAsnAspSerProSer-----AsnGluIleSerThrAspGlnLeuLysIlePheGln 167
QY 115 TTGCTGAGGAGAGTCTACGTTTATCTTTGCTCCCAACAGTGGTTCACCCGTGATC 174
Db 168 ArgMetAspGluMetSer-----AlaArgMetIle 177
QY 175 GCTCTAGATTCGGCTGTGTGTCTATTACTATAAAGATGCAAAACCCAAATATGCCAA 234
Db 178 GluMetGluGluSerPheAsnLysLeuSerAsnLysIleAlaGluGlnAsnThrMetVal 197
QY 235 ACCCAATTAGAACGACCGCGAATGGAGTTACAGATCTAATCAATGCTTAAGCGATGACA 294
Db 198 LeuAsnLeuLysGlnAspAsnTyrLysValMetAsnLysLeuAsnIleLeuLysLeu 217
QY 295 TTAGCT-----TCACATCAAGACTATGCCAAGATTGAA 327
Db 218 ValAlaGlnProSerAlaArgProSerThrAsnAsnAlaGlnAsnLysLeuAlaIleGlu 237
QY 328 -----GCTAGTTTATCATCTGTTATAGTGAAGCTGAAACAGTAAACAT 372
Db 238 LeuLeuAsnSerIleSerAlaValSerSerAlaTyrLeuGlnLysMetGlnAsnAsnGly 257
QY 373 AACCTTAATGCAACATTAGAACAACTAAATAATGGCTAAACTAATTTAGATCAGCCATC 432
Db 258 SerGlyArgGlnHisThrAlaAspLeuCysThrGlyAspSerAsnThrHisSerGlyIle 277
QY 433 AACCAA-----GCTAATACGGATTAACACGCTTTTGAT 465
Db 278 AsnGlnHisArgThrThrAsnGlyThrIleAspValAsnThrAsnThrAlaGlnLeuAsn 297
QY 466 AATGAACACCAAAATTTAGTTGAACATACAAAGCACTAAACACCTTTTAGAACCAACGT 525
Db 298 AsnGlnPheSerAsn-----AlaLeuAsnThrIleLeuProAspGln 311
QY 526 GCTACTAACCTTGAAGGTTTGTCTCATCAACTGCTTATAATCAATTCGCAATTAATTAGTG 585

```


or send an email to license@isb-sib.ch).

58	ACTCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCACAGTTG	117
QY		
Db		
13	AsnAsnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr	32
QY		
118	TC TGAGGAGAGAGTCTACGCTTTTATCTTGTGTCCTCCCAACGAGTGGGTTCACCCGATGCCGT	177
Db		
33	AsnAsnAsnAsnAsn-----ProAlaAsnAsnThrThrAsnAsnAsn	45
QY		
178	CTAGAATTCGGCTGTATGCTTATTACTAAAAAAGAGCAAAACCAATAATATGCCCAACCC	237
Db		
46	AsnSerThrGlyHisSerSerAsnThrAsnAsnAsnThrAsnAsnAsnAsnThrAsnThr	65
QY		
238	CNAITTAGAAGCAGCGCGNATGGAGTTAACAGACTTAATCAATGCTAAAGCGATGACATTA	297
Db		
56	GlyAlaSerGly-----ValAspAspPheGlnAsn-----	75
QY		
298	GCTTCACATACAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATTAGTGAAGCT	357
Db		
76	-----PhePheAspProLysProPheAspGlnAsnLeuAspSerAsnAsnAsnAsnSer	93
QY		
358	GAACAGTTACATACCTTTAATGCAACATTAGACAACATAAAAAATGGCTAAAAATAAT	417
Db		
94	AsnSerAsnAsnAsnAspAsnAsnSer-----AsnThrValAlaSerSerThrAsn	111
QY		
418	TTAGAATCA-----GCCATCAACCAAGCTAATACGGATAAAACGACT-----	459
Db		
112	PheThrSerProThrAlaValValAsnAsnAlaAlaProAlaAsnValThrGlyGlyLys	131
QY		
460	-----TTTGATATGAACACCCCAATTTAGTTGAAGCATACAAAGCACTAAAA	507
Db		
132	AlaAlaAsnPheIleGlnAsnGlnSerProGlnPheAsnSerProTyrAspSerAsnAsn	151
QY		
508	ACCACCTTTAGAACACGCTGCTACTACCTTGAGGTTTGTCTATCACTGCTTATATCAAA	567
Db		
152	SerAsn-----ThrAsnLeuAsnSerLeuSerProGlnAlaIle-----Leu	165
QY		
568	ATTGCGCAATATTAGTGGATCTATACAATAAGCTAGTAGTTTAACTAACTAACACACTA	627
Db		
166	AlaLysAsnSerIleIleAsp-----SerSerAsnLeuProLeuGlnAlaGln	181
QY		
628	GATCCACTAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAGAAT	687
Db		
182	GlnGlnLeuTyrGlyGly-----AsnAsn-----AsnAsn	189

623	ThrAsnAspLysLysSerArgAlaThrGlnGluLeuLysArgValThrGluMetLysAsn	641
370	QY -----ATAACCTTAATGCAACATTAGACAACACTAAAAATGGCTAA	411
643	SerIleGlnIleLysLeuAsnAsnLeuArgSerThrHisAspGlnAsnValLysGlnThr	662
412	ACTAATTAGAAATCAGCCATCAACCAAGCTAATACGGATAAACAGCACTTTTGATATGA	471
663	GluGlnLeuGluAlaGlnValLeuGluValAsnLysGluAsnGluThrLeuAlaGlnGln	682
472	CACCCAAATTTAGTGAAGCATACAAGCACTAAAAACCCTTTAGACAACAGTCTACT	531
683	LeuAlaValSerGluAlaAsnTyrHisAlaGlnSerLysLysLeuAsnGluLeuThrThr	702
532	AACTTGAAGCTTTGTCATCAACTGCTATATAATCAAAATTCGCAATAATTTAGTGGATCTA	591
703	AspLeuGlnGlu---SerGlnThrLysAsnAlaGluLeuLysGluGlnIleThrAsnLeu	721
592	TACAATAAGCTAGTAGTTTAATCACTAAACACTAGATCCACTAAAT	639
722	AsnSerMetThrAlaSerLeuGlnSerGln-----LeuAsnGluLysGlnGln	737
640	-----GGGGGAACGCTTTTAGTTCTTAATGATCTTAATGAG	666
738	GlnValLysGlnGluArgSerMetValAspValAsnSerLysGlnLeuGluLeuAsnGln	757
667	ATTACTACAGCTAAAT	681
758	ValThrValAlaAsnLeuGlnLysGluIleAspGlyLeuGlyLysIleSerValTyr	777
682	-----AGAAATATTAATACGTTATCACTATTAAATGAACAAGACTAAT	729
778	LeuThrLysGlnLysGluLeuAsnAspTyrGlnLysThrValGluGluGlnHisAlaGln	797
730	GCTGATGCA-----TTATCTAATAGTTTATT	759
798	LeuGlnAlaLysTyrGlnAspLeuSerAsnLysAspThrAspLeuThrAspArgGluLys	817
760	AAAGTGATTCAAAATAATGAACAA	792
818	GlnLeuGluGluArgAsnArgGlnIleGluGluGlnGluAsnLeuTyrHisGlnHisVal	837
793	GGGACTTTTCAACAGCT-----AATGTTCAACCTTCAAACTACAGTTTGTGCTTT	846
838	SerLysLeuGlnGluMetPheAspAspLeuSerGlnArgLysAlaSerPheGluLysAla	857
847	AGTGTCTGATGAACACCGTCGAATATATAATGATGACAGAGGACCGTTGGATGCTGAT	906
858	AspGlnGluLeuLysGluArgAsnIleGluTyrAlaAsn-----AsnValArg	873
907	GAACCTTCAAGTAGA-----	921
874	GluLeuSerGluArgGlnMetAsnLeuAlaMetGlyGlnLeuProGluAspAlaLysAsp	893
922	ATTCTTGCAACACGAATPATCATCAGATGTTCTTGGATTATAGTTTAGCTGGAACA	981
894	IleIleAlaLysSerAlaSerAsnThrAsp-----Thr	904
982	ACACGAGTAGTACCAATTTAGTTTGGACCACTAGTCCATCACTGGTTATTTATATTTC	1041
905	ThrThrLysGluAlaThrSerArgGlyAsnValHisGluAspThr-----	919
1042	CCTTATAAGTTGGTTAAAGCAGCT-----	1066
920	ValSerLysPheValGluThrThrValGluAsnSerAsnLeuAsnValAsnArgValLys	939
1066	-----GATGCTAATAAC	1077
940	AspAspGluGluLysThrGluArgThrGluSerAspValPheAspArgAspValProThr	959
1078	GTTGGATTACAATACAAATTAATATGGAATGTTCAACAAGTTGAGTTTGCCACTTCA	1131
960	LeuGlySerGlnSerAspSerGluAsnAlaAsnThrAsn-----AsnGly	974

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; Z35808; CAA84867.1; -;
DR EMBL; X78214; CAA35048.1; -;
DR EMBL; Z23261; CAA80797.1; -;
DR PIR; S45781; S45781;
DR SGD; S0000143; EDEL;
DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.
DR GO; GO:0005935; C:bud neck; IDA.
DR GO; GO:0005934; C:bud tip; IDA.
DR GO; GO:0006897; P:endocytosis; IMP.
DR InterPro; IPR000261; EPS15_homology.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00054; EFh; 2.
DR SMART; SM00027; EH; 3.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS50031; EH; 3.
DR PROSITE; PS50030; UBA; 1.
DR KW Hypothetical protein; Repeat.
DR FT DOMAIN 14 113 EH 1.
FT FT 135 227 EH 2.
FT FT 277 366 EH 3.
FT FT 1338 1380 UBA.
SQ SEQUENCE 1381 AA; 150783 MW; 626FD261DCBA7D99 CRC64;

Alignment Scores:
Pred. No.: 0.235 Length: 1381
Score: 140.00 Matches: 119
Percent Similarity: 28.64% Conservative: 56
Best Local Similarity: 19.48% Mismatches: 212
Query Match: 6.02% Indels: 224
DB: 1 Gaps: 22

US-09-147-052-1 (1-1371) x YBE7_YEAST (1-1381)

QY	55	ACGAATCATCTCCGAGTACCCAAATGTGACATCAAGAGAAAGTTGTTTCGAGCGTCCAG	114
Db	487	ThrProThrSerAsnPheGlyGlnSerIleLysGluGluProGluGlnGluGln	506
QY	115	TTGCTCTGAGGAAGAGCTACGTTTATCTTTGTCGCCACCA	156
Db	507	LeuArgGluSerSerAspThrPheSerAlaGlnProProValProLysHisAlaSer	526
QY	156	-----	156
Db	527	SerProValLysArgThrAlaSerThrThrLeuProGlnValProAsnPheSerValphe	546
QY	157	-----GTGGGTTTCAACCGTG	171
Db	547	SerMetProAlaGlyAlaAlaThrSerAlaAlaThrGlyAlaAlaValGlyAlaAlaVal	566
QY	172	ATCCGCTCAGAAATCGGCTGTATGCTATTACTATAAAGATCGAACCCAAATATATGGC	231
Db	567	GlyAlaAlaAlaLeuGlyAlaSerAlaPheSerArgSerSerAsnAsnAlaPheLysAsn	586
QY	232	CAA-----ACCAATTAGAACGCGCGCAATGGAGTTA	264
Db	587	GlnAspLeuPheAlaAspGlyGluAlaSerAlaGlnLeuSerAsnAlaThrThrGluMet	606
QY	265	ACAGATCTAATCAATGCTAAAGCGGATGACATTAGCTTCACTACAAGACTATGCCAAGATT	324
Db	607	AlaAsnLeuSerAsn-----GlnValAsnSerLeuSerLysGlnAlaSerIle	622
QY	325	---GAAGCTAGTTTATCATCTGCTTATAGTGAAGAGCTGAACACAGTTAAC	369

QY 1138 ACTAGTGCAGTAATACTACACTATCA-----ACTCCAGCAGTTGATGAGATTAAA 1191
 DB 975 ThrGlnSerGlyAsnGluThrAlaAsnProAsnLeuThrGluThrLeuSerAsp----- 992
 QY 1192 GTTGCTAAATCGTTTATCAGTTTAAAGATTGGCCAAACACACATGCAATTAAGTGT 1251
 DB 993 -----ArgPheAspGlyAspLeuAsnGluThrGlyLeu 1003
 QY 1252 CCAACGGGTGAGGAATATGAATAAGTTGCG----- 1284
 DB 1004 ProArgSerGlnSerLeuThrSerValAlaAsnAsnAlaProGlnSerValArgAsp 1023
 QY 1285 -----CCAAATGATGGCAACATTATCTTAGC 1311
 DB 1024 AspValGluLeuProGluThrLeuGluArgAspThrIleAsnAsnThrAlaAsnArg 1043
 QY 1312 TCAATGAAATTAATGCTGATAAGATCCCGGG 1344
 DB 1044 AspAsnThrGlyAsnLeuSerHisIleProgly 1054

RESULT 10
 PSTL_YEAST
 AC Q12355; STANDARD; PRT: 444 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Protoplast secreted protein 1 precursor
 GS PSTL OR YD055W OR D4214 OR YD9609.09
 OS Saccharomyces cerevisiae (Baker's yeast)
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96381250; PubMed=8789263;
 RA Brandt P., Ramlow S., Otto B., Bloecker H.;
 RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
 of Saccharomyces cerevisiae chromosome IV.";
 RL Yeast 12:85-90(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION
 RX MEDLINE=99251092; PubMed=10234784;
 RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
 RT "Two-dimensional analysis of proteins secreted by Saccharomyces
 cerevisiae regenerating protoplasts: a novel approach to study the
 cell wall.";
 RL Yeast 15:459-472(1999).
 RN [4]
 RP GPI-ANCHOR.
 RX MEDLINE=20469049; PubMed=11016834;
 RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
 RT "Up-regulation of genes encoding glycosylphosphatidylinositol
 (GPI)-attached proteins in response to cell wall damage caused by
 disruption of FKS1 in Saccharomyces cerevisiae.";
 RL Mol. Gen. Genet. 264:64-74(2000).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
 CC SECRETED BY REGENERATING PROTOPLASTS.
 CC -1- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC
 DR EMBL: X84162; CRA58971.1; -
 DR EMBL: Z74351; CRA98873.1; -
 DR EMBL: Z49209; CRA89084.1; -
 DR PIR: S54039; S54039.
 DR COMPLUYEAST-2DPAGE; Q12355; -
 DR SGD: S0002462; PSTL.
 DR GO: GO:009277; C:cell wall (sensu Fungi); IDA.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR Pfam: PF01030; Recep_L_domain; 1.
 KW Glycoprotein; Membrane; GPI-anchor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 ? PROTOPLAST SECRETED PROTEIN 1.
 FT PROPEP ? 444 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 356 416 SER-RICH.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 444 AA; 45776 MW; 230F60CACA5921A4 CRC64;

Alignment Scores:
 Pred. No.: 0.338 Length: 444
 Score: 138.00 Matches: 90
 Percent Similarity: 37.37% Conservative: 95
 Best Local Similarity: 18.18% Mismatches: 184
 Query Match: 5.93% Indels: 126
 DB: 1 Gaps: 20

US-09-147-052-1 (1-1371) x PSTL_YEAST (1-444)

QY 1 ATGCACATTTTAGCGGGAATTCATATTTTCCTTATAGTTATTTCTATATGTTATTTCTAGCAAC 60
 DB 3 LeuHisSerLeuIleAlaSerThrAlaLeuLeuIleThrSerAlaLeuAlaAlaThrSer 22
 QY 61 TCATCT-----CCGAGTACCCAAATGTCACATCAAGACAGTCTTTCGACGCTC 111
 DB 23 SerSerSerSerIleProSerSerCysThrIleSerSerHisAlaThrAlaAla---- 41
 QY 112 CAGTTCTCTGAGGAGAGTCTACGTTTATCTTTTGTCCCCACGAGTGGTTCACCCGNG 171
 DB 42 -----GlnSerAspLeuAspLysTyrSerArgCysAspThrLeuValGlyAsnLeuThr 59
 QY 172 ATC-----CGTCTAGATTCGGCTGTATGCTTATTTACTATAAAGATGCAACCCAAAT 225
 DB 60 IleGlyGlyGlyLeuIleThrGlyAlaLeuAlaAsnValLysGlu-----Ile 75
 QY 226 AATGGCCAA-----ACCAATTAGACGACCGCGAATG 258
 DB 76 AsnGlySerLeuThrIlePheAsnAlaThrAsnLeuThrSerPheAlaAlaAspSerLeu 95
 QY 259 GAG-----TTAACAGATCTAATCAATGCTAAAGCGATGACATTAAGTTCACCTACAAGACTAT 315
 DB 96 GluSerIleThrAspSerLeuAsnLeuGlnSerLeuThrIleLeuThr----- 111
 QY 316 GCCAAGATTGAGCTAGTTTATTCATCTGCTTATATGCTATATGAGCTGAACAGTTAAACAATAC 375
 DB 112 -----SerAlaSerPheGlySerLeuGlnSerValAspSerIle 124
 QY 376 CTTAATGCAACATTAGAACAACTAAATAAATGGCTAAACTAATTTTAAAGATGACCCATCAAC 435
 DB 125 LysLeuIleThrLeuProAlaIleSerPheThrSerAsnIleLysSerAlaAsnAsn 144

QY 436 CAAGCTAATACGATGATAAAGACGCTTTTGATAATGAACACCAAAATTTAGTTGAAGCATAC 495
 Db 145 IleTyrIleSerAspThrSerLeu-----GlnSerValAspGlyPhe 158
 QY 496 AAAGCACTAAACACCACTTTAGAACACCGTGCTACTAACCTGCAAGTTGTGCATCA--- 552
 Db 159 SerAlaLeuLysLysValAsnValPheAsnValAsnAsnLysLysLeuThrSerIle 178
 QY 553 ---ACTGCTTATATCAAAATTCGAATATTTAGTGGATCTATACATAAAGCTAGTAGT 609
 Db 179 LysSerProValGluThrValSerAspSerLeuGlnPheSerPheAsnGlyAsnGlnThr 198
 QY 610 TTAATCACTAAACACAGTACCTCACTA-----AATGGGAGACGCTTTTAGATTCT 660
 Db 199 LysIle-----ThrPheAspValLeuValTrpAlaAsnAsnIleSerLeuThrAspVal 216
 QY 661 AATGAGATGACTACAGCTAATAAGAT-----ATTATAAT 696
 Db 217 HisSerValSerPheAlaAsnLeuGlnLysIleAsnSerSerLeuGlyPheIleAsnAsn 236
 QY 697 ACCTTATCACTATTATGACAAAGACATAATGCT-----GATGCATTA 741
 Db 237 SerIleSerSerLeuAsnPheThrLysLeuAsnThrIleGlyGlnThrPheSerIleVal 256
 QY 742 TCTAATAGCTTTTATTAATAA----- 762
 Db 257 SerAsnAspThrLeuLysAsnLeuSerPheSerAsnLeuSerThrIleGlyGlyAlaLeu 276
 QY 763 GTCACTCAAAATAT----- 777
 Db 277 ValValAlaAsnAsnThrGlyLeuGlnLysIleGlyLeuAspAsnLeuThrThrIle 296
 QY 778 -----GAACAAAGTTTGTAGGACTTTTACAACGCTAATGTTCAACCTTCAAACTAC 831
 Db 297 GlyGlyThrLeuGluValValGlyAsnPheThrSerLeuAsnLeu-----AspSerLeu 314
 QY 832 AGTTTGTGCTTTTGTAGTCTGATGATACACCGCTCAATTAATAATATGCAACAGGACC 891
 Db 315 LysSerValLysGlyGlyAlaAspValGluSerLysSerSerAsnPheSerCysAsnAla 334
 QY 892 GTTGTGAATGGTGCATGAACCTTCAAGTAGAATCTTGCACAAACAGCAATATACAGAT 951
 Db 335 Leu-----LysAlaLeuGlnLysLysGlyGlyIleLysGly 346
 QY 952 GTTCTGTGATTTATAGTTAGTGGACAAACAGCAAGTACCAATTTAGTTTAGCAAC 1011
 Db 347 GluSerPheValCysLysAsnGlyAlaSerSerThrSerValLysLeuSerThrSer 366
 QY 1012 TATGTTCCATCACTGGTATTATTTATTTCCCTTATAAGTTGGTTAAAGCAGCTGATCT 1071
 Db 367 LysSerGlnSerSer-----GlnThrThrAlaLysValSerLysSer 380
 QY 1072 AATAAGCTTGGATTACAAATACAAATTAATAATGAAATGTTCAACAAAGTTAGTTTGGC 1131
 Db 381 SerSerLysAlaGluLysLysPheThrSerGlyAspIleLysAlaAlaSerAla 400
 QY 1132 ACTTCACTAGTGCATAATATACAGCTAATCACTCACTCACTCACTCACTGATGATGAT 1191
 Db 401 SerSerValSerSerGlyAlaSerSer----- 410
 QY 1192 GTTGTCAAAATCGTTTATCAGTTTAAAGTTTGGCCCAACAAACATCAATTAAGTTGT 1251
 Db 411 -----SerSerLysSerSerLysGlyAsnAlaAlaIleMetAla 424
 QY 1252 CCAACGGGTGAAGAAATATGAATAAGTTGCGCCCAATGATGGC 1296
 Db 425 ProIleGlyCln-----ThrThrProLeuValGly 434
 RESULT 11
 Y109_YEAST
 ID Y109_YEAST
 AC P40412;

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 99.7 kDa protein in SBL1 5' region precursor.
 GN Y1169C OR Y19402.07C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Moulé S., Odell C., Pearson D., Rajadream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX";
 RL Nature 387:84-87(1997).
 CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z46921; CAA87023.1; --
 DR PIR: S50358; S50358.
 DR SGD: S0001431; Y1169C.
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR PROSITE: PS50111; CHEMOTAXIS_TRANSDDUC_2; 1.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 995 HYPOTHETICAL PROTEIN Y1169C.
 FT DOMAIN 26 253 METHYL-ACCEPTING TRANSDUCER.
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 664 664 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;
 Alignment Scores:
 Pred. No.: 0.315 Length: 995
 Score: 138.00 Matches: 89
 Percent Similarity: 38.21% Conservative: 107
 Best Local Similarity: 17.35% Mismatches: 175
 Query Match: 5.93% Indels: 142
 DB: 1 Gaps: 19
 US-09-147-052-1 (1-1371) x Y109_YEAST (1-995)
 QY 52 GGTACGAATCATCT-----CCGAGTACCAAAATGTGACATCAAGAGAGTT----- 99
 Db 78 GlyValSerSerValGlyProSerSerSerValSerValSerValSerGln 97
 QY 100 -----GTTTCGAGCGTCCAGTTCTCTGAGGAGAGCTACGTTTATCTTGT 147
 Db 98 SerSerSerSerValSerAspValSerSerSerValSerGlnSerSer 115
 QY 148 CCCCCACCACTGGTTCACCGTGCCTAGATTCGCTAGATTCGCTATG 195
 Db 116 AlaSerAspValSerSerSerValSerGlnSerAlaSerSerThrSerAspValSer 135
 QY 196 TCTATTACTAAAGAGATCAACCAACCAATATGCGCAACCAATAGAGCGCGCA 255
 Db 136 SerValSerGlnSerSerSerAlaSerAspValSerSerValSerGlnSerSer 155
 QY 256 ATGGAGTTAACAGATCTAATCAATGCTAAAGCCATGACATACCTACAGACTAT 315

```

Db 156 SerSerAlaSerValSerSerValSerGlnSerAlaSerSerAlaSerasp--- 174
QY 316 GCCAAGATTGAACCTAGTTTATCATCTGCTTATAGTGAAGCTCAACAGCTTAAACATAAC 375
Db 175 -----ValSerSerValSerGlnSerAlaSerSerThrSerSer 188
QY 376 CTTAATGCAACATTAGCAACAACTAAATGGCTAAACAACTAATTAGCAATCAGCCATCAAC 435
Db 189 ValSerSerValSerGlnSerSerSerAlaSerValSerValSerSerValSer 208
QY 436 CAAGCTAATACGATAAAGCAGCTTTTGTATATGAACACCCCAATTTAGTTCAAGCATAC 495
Db 209 GlnSerSerSerAlaSerValSerSerSerValSerValSerGlnSerAlaSerThr 228
QY 496 AAAGCACTAAACCACTTTAGAACACAGCTGTACTTAACCTTTGAAGTTTGTCACTCACT 555
Db 229 SerAspValSerValSerValSerGlnSerAlaSerSerThrSerGlyValSerSer 248
QY 556 GCTTATAATCAATTCGCAATATTTAGTGGATCTATCAATAAAGCTAGTAGTTTATA 615
Db 249 GlySerGlnSerValSerSer-----AlaSerGlySerSerSerSerPhePro 264
QY 616 ACTAAACACTAGATCCACTAAATGGGGAACGCTTTTGTAGTTCTTAATCAGATTACTACA 675
Db 265 GlnSerThr-----SerSerAlaSerThrAla 273
QY 676 GCTAATAAGAAATTAATAATACCTTATCACTATTAATGAACAAAGCACTAATCGTAT 735
Db 274 SerGlySerAlaThrSerAsnSerLeuSerSerIleThrSerSerAlaSerSerAlaSer 293
QY 736 GCATTATCTAATAGTTTATTAATAA----- 762
Db 294 AlaThrAlaSerAsnSerLeuSerSerSerAspGlyThrIleLeuProThrThrThr 313
QY 762 ----- 762
Db 314 IleSerGlyAspLeuThrLeuThrGlyLysValIleAlaThrGluGlyValValAla 333
QY 763 -----GTGATTCACAAATATGAACAACTTTT-----GTAGGG 795
Db 334 AlaGlyAlaLysLeuThrLeuLeuAspGlyAspLysThrPheSerAlaAspLys 353
QY 796 ACTTTTCAACAGCGTAATGTTCAACCTTCAACCTTCAACAGCTTTTGTGCTTTAGTGTAT 855
Db 354 ValTyrGlyAspLeuLeuValLysLysSerLysGluThrTyrProGlyThrGluPheAsp 373
QY 856 GTAACACCGGCAATTATTAATATGCAAGAGGACCGTTTGGATGGATGA----- 909
Db 374 IleSerGlyGluAsnPheAspValThrGlyAsn-----PheAsnAlaGluGluSerAla 391
QY 910 ---CCTTCAAGTAGAATCTTTCGAAACACAGAAATAGTATACAGATGTTCTTGGATTAT 966
Db 392 AlaThrSerAlaSerIleTyrSerPheThrProSerSerPheAsnSerGlyAspIle 411
QY 967 AGTTTAGCT---GGAACAAACAGCACTACCAATTTAGTTTTCAGCACTATGTCCTCA 1023
Db 412 SerLeuSerLeuSerLysSerLysGlyGluValThrPheSerProTyrSerAsnSer 431
QY 1024 ACTGGTTATTATATTTCCCTTATAGTTGTTTAAAGCAGCTGATGCTTAATAAGTTGGA 1083
Db 432 GlyAlaPheSerPhe-----SerAsnAlaIle--- 440
QY 1084 TTCAATAACAATTAATAATGGAATGTTTCAACAAGTTGATGTTGCCACTTCACT--- 1140
Db 441 -----LeuAsnGlyGlySerValSerGlyLeuGlnArgAspAspThrGlu 456
QY 1141 ---AGTCAATAATACTACAGCTTAATCCAACTCCAGCAGTTGTATGATGAGATTAAAGTTGCT 1197
Db 457 GlySerValAsnAsnGlyGluIleAsn----- 465
QY 1198 AAATCGTTTATCAGGTTTAAAGTTTGGCCAAACACATCGAATTTAGTTTCCAAAG 1257
Db 466 -----LeuAspAsnGlySerThrTyrValIleValGluProValSer 479

```

```

QY 1258 GGTGAAGGAATATGATAAAGTTCGCCAATGATTGGCAACATTATCTTT----- 1308
Db 480 GlyLysGlyThrValAsnIleIleSer-----GlyAsnLeuTyrLeuHisTyrPro 496
QY 1309 -----AGCTCA 1314
Db 497 AspThrPheThrGlyGlnThrValValPheLysGlyGlyValLeuAlaValAspPro 516
QY 1315 ATGAAAATAATCTGCTGATAAGATCCCC-----GGGTAC 1347
Db 517 ThrGluThrAsnAlaThrProIleProValValGlyTyr 529

```

RESULT 12

```

YKK1_CAEEL
ID YKK1_CAEEL STANDARD; PRT; 1010 AA.
AC P34278;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C02F5.1 in chromosome III.
GN C02F5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roodra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL: L14745; AAA27916.2; -.
DR WormPep: C02F5.1; CE02450.
KW Hypothetical protein.
SQ SEQUENCE 1010 AA; 113231 MW; 15E19DD975824D94 CRC64;

```

Alignment Scores:

```

Pred. No.: 0.315 Length: 1010
Score: 138.00 Matches: 107
Percent Similarity: 40.83% Conservative: 100
Best Local Similarity: 21.10% Mismatches: 164
Query Match: 5.93% Indels: 136
DB: 1 Gaps: 28

```

US-09-147-052-1 (1-1371) x YKK1_CAEEL (1-1010)

```

QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAAAGTTGTTTCGAGCGTCCAGTTGCT 120
Db 363 AsnSerProLysHisValSerLysSerMetAspLeuGluLysThrIleGlu---Ala 391
QY 121 GAGGAGAGTCTACGTTTATCTTTGTCGCCCAACAGTGGGTTCAACCGTGATCCGCTCA 180

```

```

Db      382 AlaAspLysSerThrLysTyr-----ProSerIleAlaAspGluValGluAspLeu 398
QY      181 GAATTCGGCTGATCTATTACTATAAAGAGT-----GCAACCCCAATAAT 228
Db      399 AspMetAsp---MetAspIleThrGluGlnProCysGluAlaGlyAsnGlnAsn 417
QY      229 GCCCAACCCCAATTAGAACGACG---CGAATGGAGTTAACAGATCTAATCAATGCTAAA 285
Db      418 AspGlyLeuGlnGluGlnGluAspLeuMetAspIleSerValIleArgAspSerPro 437
QY      286 CGATG-----ACATTAGCTTCACATACAAAGACTATGCCAAGATTGAAGCTAGTTATCA 339
Db      438 AlavalAsnAspThrMetAlaValPheGlnSerProAlaArgValIleGlyAlaAsn 457
QY      340 TCTGCTTATAGTGAAGCTGAA----- 360
Db      458 AsnSerIleIleAspSerGlnLysSerIleValPheGlyAspGluMetSerIleAspGlu 477
QY      361 -----ACAGTTAACAAAT 372
Db      478 ThrGlnAsnAspGlyThrLeuThrLeuProLysSerAsnValGluValThrThrAsn 497
QY      373 AACCTTAATGCAACATTAGAACAACTAAATAATGGCTAAACCTAATTTAGAATCAGCCATC 432
Db      498 AspValThrSerLeuGluArgGlnGluAsnAlaSerGluAsnValSerMetIle 517
QY      433 AACCAAGCTAATACG-----GATAAAGCACTTTTGTATTAATGAACACCCAAT 480
Db      518 AsnGluSerSerValHisSerGluIleAspLysLysSerPhe-----Met 532
QY      481 TTAGTTGAAGCATACAAGCA---CTAAACACCACTTTAGAACACGCTGCTACTAACCTT 537
Db      533 LeuIleGluGluGluArgAlaPheMetHisSerMetIleAspValAlaGlnLysLeu 552
QY      538 GAA-----GGTTTCTCATCACTGCTTATATCAATCAATTCGCAATAATTTAGTGCATCTA 591
Db      553 GluAspAspGlySerSerLysThrPro-----ValIleLeu 564
QY      592 TACAATAAGCTAGTGTATTAATACTAAA----- 621
Db      565 AlaSerGlnSerAlaSerLeuAlaThrLysGluProSerAlaLeuHisAsnSerSerAla 584
QY      622 ACACATAGAT-----CCACTAATGGGGAACGCTTTAGATCTTAATGAGATTACT 672
Db      585 ThrLeuAsnAsnSerMetGluLeuAspAsnAsnThrLeuLeuLysThrMetGlnIleThr 604
QY      673 ACAGCTAATAAGAAATTAATAATTAATACGTATCAACTATTAAATGAACAA-----AAG 723
Db      605 ThrCys-----GluAspIleSerMetValHisGluSerIleAlaValGlu 619
QY      724 ACTAATGCTGATGATTAATCTAATAGTTTATTAATAAAGTGATTCAAATAATGAACAA 783
Db      620 LeuAsnSerAsnLysGluGlnGluGlnPheGlyAspGluThrLeuGlnLysAsnAspThr 639
QY      784 AGTTTGTAGG-----ACTTTT-----ACA 804
Db      640 SerAsnThrGlyAlaAsnPheThrPheGlnGlyHisAsnGluThrSerGlnIleMetAsn 659
QY      805 AACGCTAATGTTCAACCTTCAACATACAGTTTGTGCT---TTTAGTCTGATGTAACA 861
Db      660 AsnValAspSerGluAlaValAsnThrSerLysIleSerThrThrSerAlaPheAsnLeu 679
QY      862 CCCGTCATATTAAATATGCA-----AGAAGGACCGTTTGGATGGTGGATGAACCTTCA 915
Db      680 SerIleAsnGlnSerIleSerLysArgArgSerLeuLeuAsnSerAlaArgGluSer 699
QY      916 AGTAGAATCTTGCACACAGATAGTATCAGATGTTCTTGATTTATAGTTTAGCT 975
Db      700 ProArgArgValAlaLeuGluAsnSer-----IleMetSerMetAsn 713
QY      976 GGAACAAC-----ACGAAGTACCAATTTAGTTTGAACATGCTGGTCCATCA 1023

```

```

Db      714 GlyGlnThrMetGluAlaLeuThrGluTyrArg-----GlnAsnLysThrMetGln 730
QY      1024 ACTGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACCTTGA 1083
Db      731 ThrSerGlnAspSerMetPro-----SerMetSerLeuAsnAspSerGly 745
QY      1084 TTACAATAACAATTAATAATGAATGTTCAACAAGTTGAGTTTGGCCACTTCAACTAGT 1143
Db      746 ArgAspIleLeuAlaMetAsnThrSerValArg-----SerProHisLeuAsn 761
QY      1144 GCAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATC 1203
Db      762 SerSerLysThrAlaAlaProGlyThrProSerLeu----- 773
QY      1204 GTTTTATCAGGTTTAAAGATTGGCCAAACACAACTGAATTAAGTTTCCACGGGTGAA 1263
Db      774 -----MetSerGlnAsn---ValGlnLeuProProSerProGln 786
QY      1264 GGAATATGATTAAGTTGCCCAATGATTTGCCAAC---ATTATCTTAGCTCAATGAA 1320
Db      787 PheGluMetProAspPheAspProAlaValValAsnValValThrSerGluAsp 806
QY      1321 AATAATGCTGATAAGATCCCC 1341
Db      807 ProSerThrGluGlnHisPro 813

RESULT 13
ID ALS1 CANAL STANDARD; PRT; 1260 AA.
AC P46590;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Agglutinin-like protein 1 precursor.
GN ALS1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11651 / B792;
RX MEDLINE=95272392; PubMed=7752895;
RA Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;
RT "Candida albicans ALS1: domains related to a Saccharomyces cerevisiae
RL Mol. Microbiol. 15:39-54(1995).
CC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC -!- SIMILARITY: TO YEAST SAG1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L25902; AAC41649.2; -
KW Cell adhesion; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 1260
FT DOMAIN 433 792
FT REPEAT 433 468
FT REPEAT 468 504
FT REPEAT 504 540
FT REPEAT 540 576
FT REPEAT 576 612
FT REPEAT 612 648
FT REPEAT 648 684
FT REPEAT 684 720
FT REPEAT 720 756
FT REPEAT 756

```

```

FT REPEAT 757 792 1-10
FT DOMAIN 983 1152 2 X 26 AA APPROXIMATE REPEATS.
FT REPEAT 983 1043 2-1.
FT REPEAT 1092 1152 2-2.
FT DOMAIN 399 404 POLY-THR.
FT DOMAIN 408 418 POLY-THR.
FT DOMAIN 450 455 POLY-THR.
FT DOMAIN 486 491 POLY-THR.
FT DOMAIN 522 527 POLY-THR.
FT DOMAIN 538 563 POLY-THR.
FT DOMAIN 594 599 POLY-THR.
FT DOMAIN 630 635 POLY-THR.
FT DOMAIN 666 671 POLY-THR.
FT DOMAIN 702 707 POLY-THR.
FT DOMAIN 738 743 POLY-THR.
FT DOMAIN 774 779 POLY-THR.
FT DOMAIN 874 877 POLY-SER.
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 687 687 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 723 723 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 886 886 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 918 918 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;

Alignment Scores:
Pred. No.: 0.309 Length: 1260
Score: 138.00 Matches: 103
Percent Similarity: 36.70% Conservative: 75
Best Local Similarity: 21.24% Mismatches: 189
Query Match: 5.93% Indels: 118
DB: 1 Gaps: 21

US-09-147-052-1 (1-1371) x ALSI_CANAL (1-1260)
QY 52 GGTACGACATCA-----TCGCCAGTACCCAAAATGTGACATCAAGAGAGTT 99
DB 748 GlyThrAspThrValIleIleArgGluProAsnProThrValThrThrGluTyr 767
QY 100 GTTTCGACGCTCCAGTTGCTGTGAGGAGAGCTACGTTTATCTTTGTCGCCACACGAGTG 159
DB 768 TrpSerGln-----SerPheAlaThrThrThrValThrAlaProGlyGly 784
QY 160 GGTTCACCGCTGATCCGCTAGATTGGGCTGATGTCTATTACTATAAAGATGCAAAAC 219
DB 785 ThrAspThrValIleIleIleTyrGlu-----SerMetSerSerSerLysIleSerThrSer 802
QY 220 CCAAA-----SerMetSerSerSerLysIleSerThrSer 802
DB 803 SerAsnAspIleThrSerIleIleProSerPheSerArgProHisTyrValAsnSerThr 822
QY 235 ACCAA-----TTAGAAGCAGCGCGAATGAGTAAACAGATCAATCAATGCT 282
DB 823 ThrSerAspLeuSerThrPheGluSerSerMetMetAsnThrProThrSerIleSerSer 842
QY 283 AAAGCATGATACATTAGCTTCACTACAGACTATGCCAAGATTGAAGCTAGTTTATCATCT 342
DB 843 AspGlyMetLeuLeuSerSer-----ThrThr 851
QY 343 GCTTATAGTGAAGCTGAAACAGTTTAAACAATAACCTTAATGCAACATTAGAACAACTAA 402
DB 852 LeuValThrGluSerGluThrThrThrGluSerIleCysSerAspGlyLysGlyCys---- 870
QY 403 ATGGCTAAACAACTAATTTAGATCAGCCATCAACCAAGCTAATACGATAAAACGACTTTT 462
DB 871 -----SerArgLeuSerSerSerSerGlyIleValThrAsnPro 883
QY 463 GATAATGAACACCAAAATTTAGTTGAGCATACAAGCACTAAACCAACCTTTAGAACAA 522

```

```

DB 884 AspSerAsnGluSerSerIleValThrSerThrValProThrAlaSerThrMetSerAsp 903
QY 523 CGTGCTACTAACCTTGAAGTTTCTCATCACTGCTTATATCAATCAATCGCAATAATTTA 582
DB 904 SerLeuSerSerThrAspGlyIleSerAlaThrSerSerAspAsnValSerLysSerGly 923
QY 583 GTGCATCTATACATAAAGCTAGT---AGTTTAACTAAACACACTAGATCCACTAAAT 639
DB 924 ValSerValThrThrGluThrSerValThrThrIleGlnThrThrProAsnProLeuSer 943
QY 640 GGGGA-----ACGCTTTTATAGATTCTTAATGAG 666
DB 944 SerSerValThrSerLeuThrGlnLeuSerSerIleProSerValSerGluSerGluSer 963
QY 667 ATTACTACAGCTAATAAGATATTAATAATCACTATTAATCACTAATAATCACTAATAAG 726
DB 964 LysValThrPheThrSerAsnGlyAspAsnGlnSerGlyThrHisAspSerGlnSerThr 983
QY 727 AATGCTGAT-----GCATTATCTAATAGTTTATTATAAAGATGATCAA 771
DB 984 SerThrGluIleGluIleValThrThrSerSerThrLysValLeuProProValValSer 1003
QY 772 AATAATGAACAAAGTTTGTAGGACTTTTACAAAGCTAATGTTCAACCTTCAAACTAC 831
DB 1004 SerAsn---ThrAspLeuThrSerGluProProThrAsnThrArgGluGlnProThrLeu 1022
QY 832 AGTTTTTGT-----GCTTTTGTGCTGATGTAACACCCGTCATTAATAAATGCAAGA 885
DB 1023 SerThrThrSerAsnSerIleThrGluAspIleThrThr-----SerGln 1037
QY 886 AGGACCGTTTGGAAATGATGAACCTTCAAGTAGAATCTTCTCAACACAGATATATC 945
DB 1038 ProThrGlyAspAsnGlyAspAsnThrSerSer-----ThrAsnProVal 1052
QY 946 ACAGATGTTTCTTGGATTATAGTTTGTAGTGGACAAACACGAGTACCAATTTTACTTT 1005
DB 1053 ProThrValAla---ThrSerThrLeuAlaSerAlaSerGluGluAspAsnLysSerGly 1071
QY 1006 AGCAACTATGTCCTCAACTGCTGTTATTTATTTATTTCCCTTTAAGTTGGTTAAAGCAGCT 1065
DB 1072 SerHisGluSerAlaSerThrSer-----LeuLysProSer 1083
QY 1066 GATGCTAATAAGCTTGGATTACAAATACAAATTAATAATGAAATGTTCAACAAGTTGAG 1125
DB 1084 MetGlyGluAsnSerGlyLeu----- 1090
QY 1126 TTTGCCACTCAACTAGTGCATAATATACACTACAGCTAATCCAACT-----CCA 1173
DB 1091 ---ThrThrSerThrGluIleGluAlaThrThrThrSerProThrGluAlaProSerPro 1109
QY 1174 GCAGTTGATGAGATTAAAGTTCTCTAAATTCGTTTATCAGGTTTAAAGATTTCGCCAAAAC 1233
DB 1110 AlaValSerSerGlyThrAspValThrThrGluProThrAspThrArgGluGlnProThr 1129
QY 1234 ACAATCGAATTAAAGTTGTTCCACGCGGTGAGGAATATGATAAAGTTGCGCAATGATT 1293
DB 1130 ThrLeuSer-----ThrThrSerLysThrAsnSerSerGluSerValAlaThrThr--- 1145
QY 1294 GCAACATTTATCTTATGCTCAATGAAATATATGCTGATAGATCCCGGGTACGCTGCA 1353
DB 1146 -----GlnAlaThrAsnGluAsn-----GlyGlyLysSer 1155
QY 1354 CCCGCTACATTTTAA 1368
DB 1156 ProSerThrAspLeu 1160

```

RESULT 14

```

ID YM96_YEAST
AC Q04893;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```


DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DT Hypothetical 113.1 kDa protein in PRE5-FET4 intergenic region.
 GN YMR317W OR YMR924.09.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagels K., Lye G., Moute S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 XIII.";
 RL Nature 387:90-93(1997).
 CC -1- DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z54141; CAA90835.1; -;
 DR SGD; S0004936; YMR317W.
 KW Hypothetical protein; Repeat.
 SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Alignment Scores:

Pred. No.:	0.356	Length:	1140
Score:	137.00	Matches:	99
Percent Similarity:	37.31%	Conservative:	76
Best Local Similarity:	21.11%	Mismatches:	188
Query Match:	5.89%	Indels:	106
DB:	1	Gaps:	20

US-09-147-052-1 (1-1371) x YMR6_YEAST (1-1140)

QY	55	ACGACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTTGTTCCGAGCTCCAG	114
DB	412	ThrSerSerValSerSerGluLeuSerSerThrLysSerSerValMetSerSerGlu	431
QY	115	TGCTGTGAGGAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGTTCACCGTGATC	174
DB	432	ValSerSerAlaThrSerSerLeuValSerSerGluAlaPro-----SerAlaLeuSer	449
QY	175	CGTCTAGAAATTCGGCTGTATCTCTATTACTATAAAGATGCAACCCAAATTAATGGCAA	234
DB	450	SerLeuAlaSerSerArgLeuPheSerSerLys-----AsnThrSerValThrSer	466
QY	235	ACCAATTTAGAGAGCCGAGTGGAGTTACAGATCTAATCAATGCTAAAGCGATGACA	294
DB	467	ThrLeuValAlaThrGluAlaSerSerValThrSerSerLeuValArgProSerSerGluThr	486
QY	295	TGAGTCTCACTACAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTATAGTCAA	354
DB	487	LeuAlaSer-----AsnSerIleIleGluSerSerLeuSerThrGlyTyrAsn---	502
QY	355	GCTGAACAGTTAAACATTAACCTTAATGCAACATTTAGAACAACTAAATAATGGCTAAAC	414
DB	503	---SerThrValSerThrThrThrSerAlaAlaSerSerThrLeuGlySerLysValSer	521
QY	415	AATTAGATCAGCCATCAACAGCTAATACGGATAAAGACGCTTTTGATTAATGAACAC	474
DB	522	SerSerAsnArgMetAlaThrSerLysThrSerSerThrSerSerAspLeuSerLys	541
QY	475	CCAAATTTAGTTGAGCATCAACAGCACTAAACCACTTTAGAACACAGCTGCTACTAAC	534

RESULT 15

YEF3_YEAST	STANDARD;	PRT;	956 AA.
ID YEF3_YEAST			
AC P32618;			
DT 01-OCT-1993 (Rel. 27, Created)			
DT 01-OCT-1993 (Rel. 27, Last sequence update)			
DT 15-SEP-2003 (Rel. 42, Last annotation update)			

DB	542	SerSerValIlePheGlyAsnSerSerThrValThrThrSerProSerAlaSerIleSer	561
QY	535	CTT-----GAAGTTTGTTCATCACTCACTCACTTAATAATCAA	567
DB	562	LeuThrAlaSerProLeuProSerValTrpSerAspIleThrSerSerGluAlaSerSer	581
QY	568	ATTGCAATAATTTAGTGGATCTATACAAATAAGCTAGTAGTTTAACTAAACACACTA	627
DB	582	IleSerSerAsnLeu-----AlaSerSerSerAlaProSerAsp---	594
QY	628	GATCCACTAAATGGGGGAAACGCTTTTACATCTATAGATGATTACACCTAATAAGAA	687
DB	595	-----AsnAsnSerThrIleAlaSerAlaSerLeuIleValThrLysThrLysAsn	611
QY	688	-----ATTAATAATACGTTTACACTACTTAAATGAACAAAGACT-----AAT	729
DB	612	SerValValSerSerIleValSerSerIleThrSerSerGluThrThrAsnGluSerAsn	631
QY	730	GCTGATGATATCTATCTATAGTTTATTAATAAAGTGATTCAAAATAATGAACAAAGTTT	789
DB	632	LeuAlaThrSerSerThrSerLeuLeuSerAsnLysAlaThrAlaArgSerLeuSerThr	651
QY	790	GTAGGACTTTTACAAACGCTAACTGTTCAACCTTCAAACTACAGTTTGT---GCTTTT	846
DB	652	SerAsnAlaThrSerAlaSerAsnValProThrGlyThrPheSerSerMetSerHis	671
QY	847	AGTGTGATGTAAACACCCGCTCAATTATAATATGCAAGAAGGACCGTTTGGAAATGGT	906
DB	672	ThrSerValIleThrPro-----GlyPhe	679
QY	907	GAACCTTCAAGTAGAATTTCTTGCACAAACAGAAATAGTATCAGAGATGTTCTTGG	966
DB	680	SerThrSerSerAlaSerLeuAlaIleAsnSerThrValValSer-----Ser	695
QY	967	AGTTTAGCTGGAACAAACACGAGTACCAATTTACTTTTACCACTATGGTCCATCACT	1026
DB	696	SerLeuAlaGly-----TyrSerPheSerThrProGluSerSerProThrThr	711
QY	1027	GTTTATTATATTTCCCTTAAAGTTGGTTAAAGCAGCTGATGCTTAATAGCTTGGATTA	1086
DB	712	Ser-----ThrLeuValThrSerGluAlaProSerThrValSer---	724
QY	1087	CAATACAATTAATAATGGAATGTTCAACAAGTTGAGTTGGCCTTCACTCACTAGTGA	1146
DB	725	-----SerMetThrThrSerAla	730
QY	1147	-----AATACTACAGCTAATCAACTCCAGAGTTGATGAGATTAAGTT	1194
DB	731	PropheIleAsnAsnSerThrSerAlaArgProSerProSerThrAlaSerPhe	748
QY	1195	GCTAAATCGTTTATCAGGTTTAAAGATTGGCCAAACACAACTCAATCAATTA	1251
DB	749	-----IleThrGluSerThrSerSerIleSerSerVal	759
QY	1252	CCACGGGTGAAGGAAATATGAATAA-----GTT	1281
DB	760	ProLeuAlaSerGlyAspValThrSerSerLeuAlaIleHisAsnLeuThrThrPheSer	779
QY	1282	CGGCCAATGATTGGCAACATTTCTTACCTAGCTCAAAATGAAATAATGCTGATAAGATC	1338
DB	780	AlaProSerThrSerSerAlaGlnLeuValSerLysSerThrThrSerSerSerIleLeu	799
QY	1339	-----CCCGGTACCGTCGACCGGT	1359
DB	800	ValThrProArgIleAspArgSerGly	808

D	b		216	PheSerIleAspLeuLysLysIle-----LeuIleCysAlaGlnGluAspLeu	232
Q	y		358	GAAACAGTTTAACTAATACCTTAATGCACAATTAGAACAACATAAAATGGCTAAAACCTAAT	417
D	b		233	HisaspValLeuSerGlnGlnThrSerLeuLeuGlnAspPheGlnGluSerLysLeuGlu	252
Q	y		418	TTGAATCAAGCCCATCAACCAAGCTAATACGGATAAAACGACTTTTCATATGAATGAACACC	477
D	b		253	LeuGluLeuGluLeuAsnLeu-----LysThrHisTrpSerHisGlu-----	267
Q	y		478	AATTTAGTTGAAGCATAACAAGCAGCTAAAAACACCTTTAGAACACGCTGTACTAAC---	534
D	b		268	----IleAspLeuArgLysSerLysSerAsnIleLysSerLeuGluAsnSerLys	285
Q	y		535	-----CTTGAAGGTGG-----TCATCAACTGCT	558
D	b		286	LeuLeuThrAspLeuLysIleGluLysLeuAsnLysLysIleAspLysSerLysGluLys	305
Q	y		559	TATAAATCAAATTCGCAATAATTTAGTTGGATCTATACAATAAAGCTAGTAGTTTAATACT	618
D	b		306	IleSerLysMetArgAsnAspMetGlnLysTrpSerGlnGluAspThrGluLeuLeuSer	325
Q	y		619	AAAACACTAGATCCACTAAATGGGGGAAGCGTTTTAGATTCTTAATGAGATTACTACAGCT	678
D	b		326	LysaspThrIleLysGluLysTyrrPheLysLeuLeuAsnGluSerAsnAlaSerValAla	345
Q	y		679	AATAAGAAAT-----ATTATATACGTTATCACTATTAATGAACAAAAG	723
D	b		346	AsnIleAsnLysGluLeuGluSerLeuGlnAsnGluIleSerLysMetGluGluSerAsn	365
Q	y		724	ACTAATGCTGATGATTAATCTAATATCTTTTATAAAAAAGTCATCAAAATAATGAACNA	783
D	b		366	LysArgLeuAsnAlaSerLysSerLysSerLeuIleThrSerIleVal-----	380
Q	y		784	AGTTTGTAGGAGCTTTTACAACGCTAATGTTCACACCTTCAACCTACACAGTTTGTGCT	843
D	b		381	-----ValAsnAlaAsnValGlu-----AsnAspLysProIleAla	392
Q	y		844	TTTAGTGCTGATGTAACCCGCTCAATTAATATATGCAAGAGACACGCTTGGAAAT---	900
D	b		393	---SerGlyGluLeuSerAlaValLeuLysLysLeuAsnAspPheThrLeuGluLysAsn	411
Q	y		901	-----GGTGATGAACCTTCAAGTAGAATTTCTTGCAAAACACAGTAATAGT	942
D	b		412	GlyPheLeuSerAsnAlaGlyGluPheLeuSerLysLeuAsnAlaAspSerSerLeu	431
Q	y		943	ATC-----ACAGATGTTCTCTGG-----	960
D	b		432	IleLysMetIleLysGlnGluLeuSerIleAspGlnGluLeuGluAlaAsnTriLysLeu	451
Q	y		961	-----ATTATATAGTTTAGCTGGCAACAAACACAGAGTAC---CAATTTAGT	1002
D	b		452	GlnArgSerAsnLeuLysLysIleSerAlaLeuGluAsnGlnPheAsnGluMetSer	471
Q	y		1003	TTTAGCACTATPGTTCATCACTCGTTTATTATATTTCCCTTATAAG-----	1050
D	b		472	LeuAsnAsnArgAsnLeuLysThrLysLeuMetValGlnProTyrrLysAsnAsnGlyAsp	491
Q	y		1051	TTGGTTTAAGCAGCTGATGCTTAATAACGTT-----GGATTA	1086
D	b		492	SerLeuAlaAlaThrAsnSerAsnAsnSerAlaGluLysAsnArgSerSerGlySerIle	511
Q	y		1087	CAATACAAATTAATAAT--GGAATGTTCAACAGTTGAGTTTGCC	1131
D	b		512	GlnLeuProLeuSerAsnAsnMetSerArgThrGlySerIleAspLeuIleSer---Asn	530
Q	y		1132	ACTTCACTAGTGCAAAATAACTACAGCT-----AAT	1164
D	b		531	AsnAsnLysSerIleAsnAsnSerAsnAlaAspSerAlaProProLeuArgLeuHicAsn	550
Q	y		1165	CCA-----ACTCCAGCAGTTGATCAGATTAAAGTTGCTAAATCGTTTATCATCAGGT	1215
D	b		551	ProValSerTyrrProSerAsnGlnProIleGlnProSerSerSerLeuLeuSerGln	570

QY 1216 TTAAGATTGGCCAAACACACATCGAATTAACTGTTCACCGGTGAAGGAATATGAAT 1275
Db ||| |||::||| |||
571 LeuThr-----GlnAspThrAspAsnArgSer----- 579
QY 1276 AAAGTTGGCCCAATGATTGGCAACATTTATCTTAGCTCAAATGAAAAATAATGCTGATAAG 1335
Db |||::: ||| :::::||||||| ||| :::
580 -----MetLeuSerAsn---HisIleSerSerAsnAsnGluAsnLysGlnGln 594
QY 1336 ATCCCCGGGTAC 1347
Db |||
595 ProSerSerTyr 598

Search completed: October 8, 2003, 17:55:24
Job time : 54.0964 secs

10 11 43 05 2003

THIS PAGE BLANK (USPTO)

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	832	35.8	702	2	S48754	major surface prot	
2	754	32.4	702	2	S48753	major surface prot	
3	749	32.2	650	2	S48752	major surface prot	
4	745.5	32.1	649	2	S48752	major surface prot	
5	715.5	30.8	647	2	A43218	hemagglutinin homo	
6	578	24.8	386	2	S48755	major surface prot	
7	314	13.5	865	1	VGEBB	glycoprotein B pre	
8	189.5	8.1	320	2	S53560	major surface prot	
9	177	7.6	6713	2	B89921	hypothetical prote	
10	173.5	7.5	933	2	S41539	fibrinogen-binding	
11	173	7.4	1302	1	JC5009	surface-located me	
12	166.5	7.2	3890	2	C89921	hypothetical prote	
13	166	7.2	4688	2	F82852	hypothetical prote	
14	165	7.1	661	2	AG2422	hypothetical prote	

Db 246 AspLysPhe---AsnGluPheGluLysAsnProLeuSerLysGluLysLeuLysSerThr 264
 QY 793 GGGACTTTTACAAAGCGTAAGTTCAACCTTCAACACTACAGTTTGGTTCCTTTAGTGCT 852
 Db 265 SerAspThrAlaHisAsnGlnGluGlnProAlaAsnTrpSerPheAlaAlaTyrSerVal 284
 QY 853 GATGTAAACACCC-----GTCAATTATATAATATGCAAGAGGACCGCTTTGG 897
 Db 285 AspLeuThrSerAsnSerGlnAsnLeuProAsnTrpAsnPheAlaGlnArgLysValTrp 304
 QY 898 AATGGTGAT-----GAACCTTCAAGTAGAATCTTGCAACACACGAAT---AGTATCAC 948
 Db 305 ThrSerGluAsnGlnGlnProGlyLysThrAlaLeuValSerSerProValSerAlaThr 324
 QY 949 GATGTTCTTGATTTAGTTAGCTGCGACCAACACAGAGTACCAATTTAGTTTACG 1008
 Db 325 AspValSerTrpIleTyrSerLeuAlaGlyGluGlyThrLysTyrThrLeuThrPheGlu 344
 QY 1009 AACTATGTCATCAACCTGGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGAT 1068
 Db 345 TyrTyrGlyProAspAsnAlaPheLeuTyrLeuProTyrLysLeuValLysAlaAlaAsp 364
 QY 1069 GCTATAACGTTGGATTACATAACAAATTAATATGGAATGTTCAACAAGTTCGAGTTT 1128
 Db 365 SerSerSerValAlaLeuGlnTyrSerLeuAsnLysThrSerSerLysLeuIleAsnPhe 384
 QY 1129 -----GCCACT 1134
 Db 385 LysProAlaGluThrValSerThrAsnThrAspGlnSerGluAsnGluValAlaThrThr 404
 QY 1135 TCAACT-----AGTCAAA 1149
 Db 405 SerThrThrGluAlaArgSerSerTyrLysValLeuValAlaAspGluAlaThrSer 424
 QY 1150 AATACACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTT 1209
 Db 425 AsnAsnGluMetAsnHisThrProThrValSerAspIleAsnIleAlaLysValThrLeu 444
 QY 1210 TCAGTTTAAAGTTGGCCAAACACAACTCGAATTAAGTTGTTCCAAACGGGTGAAGAAAT 1269
 Db 445 SerGlyLeuThrPheGlyGluAlaThrIleGluPheSerValProGluGly----- 461
 QY 1270 ATGAATAAGTTGCCCAATGATTCGCAACATTTATCTAGCTCAATGAAATGCT 1329
 Db 462 -----LysValAlaProMetIleGlyAsnMetTyrLeuThrSerAsnSerGluSerGln 479
 QY 1330 GATAAGATC 1338
 Db 480 ValLysIle 482

RESULT 3
 S48751
 major surface protein (clone pMGal.1) precursor - Mycoplasma gallisepticum
 N:Alternate names: major hemagglutinin pmGA
 C:Species: Mycoplasma gallisepticum
 C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
 C:Accession: S48751; A44793
 R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
 FEBS Lett. 352, 347-352, 1994
 A:Title: The organization of the multigene family which encodes the major cell surface
 A:Reference number: S48751; MUID:95010739; PMID:7925999
 A:Accession: S48751
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-650 <MAR>
 R:Cross-references: GB:U90714; EMBL:L28423; NID:g1905869; PIDN:AAB50152.1; PID:g1905870
 R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whithear, K.G.
 Infect. Immun. 60, 3885-3891, 1992
 A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepticum
 A:Reference number: A44793; MUID:92363591; PMID:1379991
 A:Contents: S6
 A:Accession: A44793
 A>Status: preliminary

A:Molecule type: protein
 A:Residues: 26-42 <MAR>
 A>Note: sequence extracted from NCBI backbone (NCBIP:111017)
 C:Genetic:
 A:Genetic code: SGC3
 A:Start codon: GTG

Alignment Scores:
 Pred. No.: 1,92e-39 Length: 650
 Score: 749.00 Matches: 178
 Percent Similarity: 55.61% Conservative: 65
 Best Local Similarity: 40.73% Mismatches: 138
 Query Match: 32.20% Indels: 56
 Db: 2 Gaps: 12

US-09-147-052-1 (1-1371) x S48751 (1-650)

QY 157 GTGGGTTCAACCGTCATCGCTAGAAATTCGGCTGTATCTTACTATAAAAGATGCA 216
 Db 15 IIGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProSerProAla 34
 QY 217 -----AACCCAAATAAT 228
 Db 35 ProAsnProAsnProProSerAsnGlyGlyMetAsnGlyGlyAsnIleAsnProGlyAsp 54
 QY 229 GCCCAA-----ACCCAAATTAGAGCAGCGCGAATGAGTTAACA 267
 Db 55 GlyGlnGlyMetMetAsnAlaAlaGlnGluLeuAlaAlaAlaAlaArgMetGlyLeuThr 74
 QY 268 GATCTTAATCAATGCTAAACCGGATGACATTAGCTTTCACACAGACTATCCCAAGATTGAA 327
 Db 75 ThrValPheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGln 94
 QY 328 GCTAGTTTATCATCTGCTTATAGTAGAAGCTGAACAGTTAAACAATTAACCTTAATGCAACA 387
 Db 95 AsnThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerThr 114
 QY 398 TTAGAACAACTAAATGGCTAAACTAATTTAGAACTAGCCATCAACCAACTAATAGC 447
 Db 115 ThrGlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThr 134
 QY 448 GATAAAACAGCTTTTGATTAATGAACACCCAAATTTAGTTGAAGCATCAACACCACTAAAA 507
 Db 135 SerLysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysLysLeuLys 154
 QY 508 ACCACTTTAGAACAACTGCTACTTAACCTTGAAGGTTTGTCTATCAACTGCTTATATCAA 567
 Db 155 ThrThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGly 174
 QY 568 ATTCCCAATAATTTAGTGGATCTATACAATAAGCTAGTAGTTTAACTAAACACTA 627
 Db 175 IleLysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeu 194
 QY 628 GATCCACATAAATGGGGAACGCTTTAGATTTCTAATAGATTTACAGATTAACATAAGAT 687
 Db 195 GluProValGluGlyAspPro---LeuThrAlaGlyAlaValThrMetAlaAsnThrLys 213
 QY 688 ATTAATAATACGTTA-----TCACTATTATTAACAAAGACTAATGCTGATGATTA 741
 Db 214 IleValGluAlaIleLysAspGluValLeuAsnProLysLysGluAsnAlaThrLysLeu 233
 QY 742 TCTAATAGTTTATTAATAAAGTGAATCAAAATATGAACAAGTTTGTAGGACTTTT 801
 Db 234 AlaAspSerPheValLysGlnValLeuValLysGluLysIleThrGlyValGluAla 253
 QY 802 ACAAAACGCTAATGTTCAACCTTCAACCTTCAACCTTGTGCTTTTAGTCTCATGATACA 861
 Db 254 HisAsn---LysAlaGlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThr 272
 QY 862 CCCGTC-----AATTATAAATATGCAAGAGGACCGCTTTGG 897
 Db 273 GlyThrAlaAsnGlyGlnThrSerIleProAsnTrpAsnTyrAlaGlnArgThrIlePhe 292

```

QY 898 ---AATGCTGATGAACCTTCAAGTAGAATTTCTTCAACACG----- 936
Db 293 ThrAsnGlyAspGluPro-----ArgSerValSerAsnThrProValAspGlyGlnThr 310
QY 937 -----AATAGTATCACAGATGTTCTTGGATTATATAGTTAGCTGGGAACAACACGAA 990
Db 311 MetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLys 330
QY 991 TACCATTATTAGTCAACATATGGTCCATCACTCACTGTTATTTATATTTCCCTTATAAG 1050
Db 331 TyrThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 350
QY 1051 TTGGTTAAAGCAGCTGATGCTAAACGTTGGATTACAAATCAACAATTAATAATGGAAT 1110
Db 351 LeuValAsnThrSerAspGlnMetLysLeuGlyLeuGluTyrLysLeuAsnAspAla--- 369
QY 1111 GTTCAACAAGTTGAGTTGGCATTCTCACTAGTCAATATATACTACAGCTAAT---CCA 1167
Db 370 -----ThrGluProSerAlaIleThrPheGlyAsnGluGlnThrMetAsnGlyLys 386
QY 1168 ACTCCAGCAGTGTATGAGATAAGTTGCTAAATTCGTTTATCAGGTTTAAAGTTGGC 1227
Db 387 ThrProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeuIlePheGly 406
QY 1228 CAACAACAATCGAATTAAGTGTTCACCGGTGAAGGAAATATGAATAAAGTTGCGCCA 1287
Db 407 SerAsnLysIleGluPheSerValProAla-----GluLysValSerPro 421
QY 1288 ATGATTGGCAACATTTATCTAGCTCAATGAATAATATCTGATAGATC 1338
Db 422 MetIleGlyAsnMetTyrLeuSerSerProAsnAsnTrpAsnLysIle 438

RESULT 4
S48752
major surface protein (clone pMG1.2) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48752
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48752
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-649 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Alignment Scores:
Pred. No.: 3.17e-39 Length: 649
Score: 745.50 Matches: 177
Percent Similarity: 55.00% Conservative: 65
Best Local Similarity: 40.23% Mismatches: 135
Query Match: 32.05% Indels: 63
DB: 2 Gaps: 12

US-09-147-052-1 (1-1371), x S48752 (1-649)
QY 157 GTGGGTTCAACGCTGATCGCTAGAAATTCGGCTGTATGCTTACTTACTAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrProThrProAsnProThr 34
QY 211 -----GATGCAACCCCAATAATGGC 231
Db 35 ProAsnProAsnProSerGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
QY 232 CAA-----ACCCAATTAGAAGCAGCGCGAATGAGTTAACAGT 270
Db 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74

```

```

QY 271 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 330
Db 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysThrGlnAsn 94
QY 331 AGTTTATCATCTGCTTATAGTAGTGAACAGTTAAACAATTAACCTTAATGCAACATTA 390
Db 95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerThrThr 114
QY 391 GAACAACATAAATGGCTAAACATAATTTAGATCAGCCATCAACCAAGCTAATACGGAT 450
Db 115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY 451 AAAACGACCTTTGATATGAACCCCAATTTAGTTCAACCATACACAAGCACAATAAAC 510
Db 135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY 511 ACTTTAGACAACAGTGTCTACTAACCCTTGAAGGTTTGTCACTCAACTGCTTAATCAAT 570
Db 155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY 571 CGCAATAAATTTAGTGGATCTATACAATAAGCTAGTAGTTTAATACTAAACACATAGAT 630
Db 175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY 631 CCACATAAGTGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAGAATTT 690
Db 195 ProValGluGlyAspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY 691 AATAAATACGTTA-----TCACTATTATGAACAACAAAGTCTTCTAGGACTTTTACA 804
Db 214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY 745 AATAGTTTATTAAAAAGTGAATTCAAAATAATGAACAAAGTTTCTAGGACTTTTACA 804
Db 234 AspSerPheValLysGlnValLeuValLysGluLysIleThrGlyValGluGluAlaHis 253
QY 805 AAGCTTAATGTTCAACCTCAACACTACAGCTTTTGTGTTTGTCTGTCTGATTAACACCC 864
Db 254 Asn---LysAlaGlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThrGly 272
QY 865 GTC-----AATTATAATATGCAAGAGGAGGCCCTTTGG--- 897
Db 273 ThrThrThrGlyGlnThrSerIleProAsnTrpAspTyrAlaGlnArgThrIlePheThr 292
QY 898 AATGTTGATGAACCTTCAAGTAGAATTTCTGCAACACG----- 936
Db 293 AsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAspGlyGlnThrMet 310
QY 937 ---AATAGTATCACAGATGTTCTTGGATTATAGTTAGCTGGGAACAACACGAAAGTAC 993
Db 311 AlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLysThr 330
QY 994 CAATTTAGTTTACCAACTATGCTCCATCACTGCTTATTTATATTTCCCTTATAGTTG 1053
Db 331 ThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLysLeu 350
QY 1054 GTTAAAGCAGCTGATGCTTAATCAACCTTGGATTACAAATTAATAATGAATAATGTT 1113
Db 351 ValAsnThrSerAspGlnValLysLeuGlyLeuGluTyrLysLeuAsnAsp----- 367
QY 1114 CAACAAGTTGAGTTTGGCCACTTCAACTAGTACA-----AATAATACTACA 1158
Db 368 -----AlaThrLysProSerAlaIleThrPheGlySerAspGlnThrMet 382
QY 1159 GCTAATCCAACTCCACAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTTTA 1218
Db 383 AsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeu 402
QY 1219 AGATTGGCCAAAACACAATCGAATTAAGTGTTCACAGCGGTGAAGGAAATATGAATAA 1278
Db 403 AsnPheGlySerAsnLysIleGluPheSerValProAla-----GluLys 417
QY 1279 GTTGGCCCAATGATTGGCAACATTTATCTTAGCTCAATGAATAATGCTGATAAGATC 1338

```



```

|||||
418 ValSerProMetIleGlyAsnMetTyrLeuSerSerProAsnAsnTrpAsnLysIle 437
|||||

RESULT 5
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218; MUID:93162830; PMID:8432610
A:Accession: A49218
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-647 <MAR>
A:CROSS-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A>Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)
C:Genetics:
A:Genetic code: SGC3

Alignment Scores:
Pred. No.: 2,35e-37 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 2 Gaps: 13

US-09-147-052-1 (1-1371) x A49218 (1-647)
QY 157 GTGGTTCACCGTGATCGTCGTAGATTGCGCTGTCTGTCTATTACTATAAAAAA----- 210
DB 15 IleGlySerPheValMetLeuAlaAlaAlaSerCysThrThrProThrProAsnProThr 34
QY 211 -----GATGCAACCCAAATAATGGC 231
DB 35 ProAsnProAsnProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
QY 232 CAA-----ACCCAATTTAGAACGCGCGGAATGAGTGAACAGAT 270
DB 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74
QY 271 CTAATCAATGCTAAAGCGATGACATAGCTTCACTACAGACTATGCCAAGATTGAAGCT 330
DB 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGlnAsn 94
QY 331 AGTTTATCATCTGTTAGTGAAGCTGAACAGTTTACATTAACCTTAATGCAACATTA 390
DB 95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThr 114
QY 391 GAACAACATAAAATGCTTAAACTAATTTAGATCAGCCATCAACCAACCTAATACGGAT 450
DB 115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY 451 AAAAGCAGCTTTGATTAATGAACACCCAAATTTAGTTGAAGCATACAAAGCCTAAAAACC 510
DB 135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY 511 ACTTTAGAACACAGCGCTACTAACCTTGAAGCTTTCATCACTCACTGCTTATAATCAAAAT 570
DB 155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY 571 CGCAATAATTTAGTGGATCTATACAAATAAGCTAGTATTAACTAAACACACTAGAT 630
DB 175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY 631 CCACATAATGGGGGACGCTTTTATGATCTTAATGAGATTACTACAGCTAATAAGAAATAT 690
DB 195 ProValGluGlyAspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213

```

```

QY 691 AATAATAGCTTA-----TCAACTATTATTATGAACAAAAAGACTAAAT-----GCT 732
DB 214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY 733 GATGCAATATCTAATAGTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTGA 792
DB 234 AspSerLeuLeuSerSerIleValLysIleThrGlyValGluGlu----- 249
QY 793 GGGACTTTTACAAACGCTAATGTTTCAACCTTCAAACTACAGTTTGTGCTTTTGTAGTGT 852
DB 250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY 853 GATGTAAACACCGCTC-----AATATAAATATATCAACAGG 888
DB 267 TrpTyrThrGluLeuLeuLeuAspLysGlnValPheProAsnTrpAspTyrAlaGlnArg 286
QY 889 ACCGTTTGG---AATGGTGATGAACCTTCAAGTAGAATTTCTTGCAACACG----- 936
DB 287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY 937 -----AATAGTATCACAGATGTTTCTTGATTTATAGTTTATAGTTTACGGAACA 981
DB 305 GlyGlnThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY 982 AACACGAAGTACCAATTTAGTTTGAACACTATGTCCTCACTCACTGGTTATTTATATTC 1041
DB 325 GlyAlaLysTyrThrLeuGluPheThrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY 1042 CCTTATAAGTTGGTTAAAGCAGCTGATGCTATATACGTTGGATTCAATACATCAATTAAT 1101
DB 345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGlyLeuGluTyrLysLeuAsn 364
QY 1102 AATGGAATGTTTCAACAAAGTTGAGTTGCCACTTCAACTCACTGCA----- 1146
DB 365 Asp-----AlaThrLysProSerAlaIleThrPheGlySer 376
QY 1147 AATAATATACAGCTAATCCAACTCCAGCTGATGATGAGATTAAAGTGTCTAAAATCGTT 1206
DB 377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY 1207 TTATCAGGTTTAAAGATTGGCCAAACAACTGAATTAAGTGTCCACAGGTTGAAGCA 1266
DB 397 LeuAlaAsnLeuAsnPheGlySerAsnLysIleGluPheSerValProAla----- 413
QY 1267 AATATGAATAAGTTGCCCAATGATGCGCAACATTTATCTTACCTCAATGAAATATAT 1326
DB 414 -----GluLysValSerProMetIleGlyAsnMetTyrLeuSerSerSerProAsnAsn 431
QY 1327 GCTGATAGATC 1338
DB 432 TrpAsnLysIle 435

RESULT 6
S48755
major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48755
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48755
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-386 <MAR>
A:CROSS-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692
A>Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Alignment Scores:

```

QY 967 AGTTTACGTTGGACAAACACGAGTACCAATTTAGTTTAGCAACATGATGGTCCATPCA--- 1023
 Db |||||:|||| ||||| :||| ||||| |||
 332 SerLeuSerGlyMetGlyAlaLysTyrThrValThrPheAspTyrTyrGlyAlaSerAsn 351
 QY 1024 ACTGCTATTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTTAATAACGTTGGA 1083
 Db |||||:||||| ||||| ||||| |||
 352 AsnAlaTyrLeuTyrPheProTyrIlysLeuValGlnThrAsnAsp-----AsnValGly 369
 QY 1084 TTACAATAACAATTAAATTAATGGAATGTTCAACAAGTTGAGTTT 1128
 Db |||||:||||| ||||| :||| |||
 370 LeuGlnTyrValLeuAsnAsnThrThrProLysLeuValAsnPhe 384
 RESULT 7
 VGBERB
 glycoprotein B precursor - Marek's disease virus (strain RB1B)
 C:Species: Marek's disease virus
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
 C:Accession: A32402; B32402
 J. Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.
 J. Gen. Virol. 70, 1789-1804, 1989
 A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog
 A:Reference number: A32402; MUID:89293086; PMID:2544666
 A:Accession: A32402
 A:Molecule type: DNA
 A:Residues: 1-865 <ROS>
 A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837
 A:Accession: B32402
 A:Molecule type: protein
 A:Residues: 250-271;304-330 <ROS2>
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-865/Product: glycoprotein B #status predicted <SIG>
 F:709-728/Domain: transmembrane #status predicted <TN1>
 F:732-752/Domain: transmembrane #status predicted <TN2>
 F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Alignment Scores:
 Pred. No.: 2,45e-12 Length: 865
 Score: 314.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.50% Indels: 0
 DB: 1 Gaps: 0
 US-09-147-052-1 (1-1371) x VGBERB (1-865)
 QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTAGTATTCTATATGTCACGAAC 60
 Db 1 MethHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCCCGAGTACCCAAATGTGCATCAAGAGAGTGTGTTTCGAGCGTCCAGTTGTCT 120
 Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCGCCACACAGTGGTTCACCGTGATCCGTCFA 180
 Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAA 183
 Db |||
 61 Glu 61
 RESULT 8
 S51560
 major surface protein (clone pmgal.6) - Mycoplasma gallisepticum (fragment)
 C:Species: Mycoplasma gallisepticum
 C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
 C:Accession: S51560; S48757
 R:Markham, P.F.; Giew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
 FEBS Lett. 352, 347-352, 1994
 A:Title: The organisation of the multigene family which encodes the major cell surface
 A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S51560
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-320 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62415.1; PID:g535688
C:Genetics:
A:Genetic code: SGC3

Alignment Scores:
Pred. No.: 0.000147 Length: 320
Score: 189.50 Matches: 46
Percent Similarity: 60.64% Conservative: 11
Best Local Similarity: 48.94% Mismatches: 26
Query Match: 8.15% Indels: 11
DB: 2 Gaps: 3

US-09-147-052-1 (1-1371) x S51560 (1-320)

QY 1060 GCAGCTGATGCTATAAAGCTGGATTACAAATTAATAATGAAGTTCACAA 1119
Db :|||||:|||||
Db 41 SerAlaAspSerAsn-----ProThrAsnGlnGluAsnSerGlnSer 54
QY 1120 GTTGAGTTGGCAGCTTCAACTAGTGTCAATTAATACAGCTTCAACTCCAGCAGTT 1179
Db :|||
Db 55 AsnGlnAlaAlaProGluAlaSerAla-----MetAsnGluThrProThrVal 70
QY 1180 GATGAGATAAAGTTGCTAAATCGCTTTATCAGCTTTAAGATTTGGCCAAACAAATC 1239
Db :|||
Db 71 AspGlyIleAsnValAlaLysValThrLeuThrAspLeuLysPheGlySerAsnThrIle 90
QY 1240 GAATTAAGTGTGCCA---ACGGGTGAAGGAAATATGAATGAAGTTGGCCCAATGATGGC 1296
Db :|||||:|||||
Db 91 GluLeuSerValProThrThrAspGluGluGlyThrSerLysValAlaPrometIleGly 110
QY 1297 AACATTATCTAGCTCAATCAAAATTAATGCTGATAAGATC 1338
Db :|||||:|||||
Db 111 AsnMetTyIleThrSerAsnAspGlnAsnGlnArgLysIle 124

RESULT 9
B89921
Hypothetical protein ebha [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Alignment Scores:
Pred. No.: 0.000777 Length: 6713
Score: 177.00 Matches: 105
Percent Similarity: 36.64% Conservative: 76
Best Local Similarity: 21.26% Mismatches: 191
Query Match: 7.61% Indels: 122
DB: 2 Gaps: 19

US-09-147-052-1 (1-1371) x B89921 (1-6713)

QY 58 AACTCATCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTG 117
Db :|||||:|||||
Db 465 AsnAsnLeuThrSerIleAsnAlaGlnLysArgAspLeuThrThrLysIle----- 482

QY 118 TCTGGGAAGAGTCTACGTTTTATGTTTGTCCCCACCAGCGTGGTTCACACCGTGATCCGT 177
Db :|||||:|||||
Db 483 -----AspGlnAlaThr----- 486
QY 178 CTAGAATTTCGGCTGTATGCTATTACTAAAAAAGATGCAAAACCCAAATAAATGCGCAACCC 237
Db :|||||:|||||
Db 487 -----ThrValAlaGlyValGluAlaValSerAsnThrGly---Thr 499
QY 238 CAATTAGAAGCAGCCGGAATGAGTTAAACAGATTAATCAATCAATGCTAAAGCGATGATTA 297
Db :|||||:|||||
Db 500 GlnLeuAsnThrAlaMetAlaAsnLeuGlnAsnGlyIleAsnAspLysAlaAsnThrLeu 519
QY 298 GCTTCACACAGACTATGCCAAGATTGAAGTACTGTTATCATCTGCTTATAGTGAAGCT 357
Db :|||||:|||||
Db 520 AlaSer---GluAsnTyHisAspAlaAspSerAspLysLysThrAlaTyThrGlnAla 538
QY 358 GAAACA-----GTTAAACAATAACCTTAATGCAACATTAAGCAACACTAGCAACAACTAAA 402
Db :|||||:|||||
Db 539 ValThrAsnAlaGluAsnIleLeuAsnLysAsnSerGlySerAsnLeuAsp----- 555
QY 403 ATGGCTAAACCTTAATTTAGAAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTT 462
Db :|||||:|||||
Db 556 -----LysAlaAlaValGluAsnAlaLeuSerGlnValThrAsnAlaLysGlyAlaLeu 573
QY 463 GATAATGACACCCCAATTTAGTTGACGATCAACAGCNCCTAAACACCACTTTAGAACAA 522
Db :|||||:|||||
Db 574 AsnGlyAsnHis---AsnLeuGluGlnAlaLysSerAsnAlaAsnThrThrIle----- 590
QY 523 CGTGCTACTAACTGGAAGTTGTGCATCACTGCTTATAATCAATTCGCAATAATTTA 582
Db :|||||:|||||
Db 591 -----AsnGlyLeuGlnHisLeuThrAlaGlnLysAspLysLeuLysGlnVal 608
QY 593 GTGGATCTATACATAAAGCTAGT---AGTTTAATACTAAACAACTAGATCCACTAAT 639
Db :|||||:|||||
Db 609 GlnGlnAlaGlnAsnValAlaGlyValAspThrValLysSerSerAlaAsnThrLeuAsn 628
QY 640 GGG-----GGAACGCTTTTACATCTAATGAGATTACTACAGCTAATGAAGATATTAT 693
Db :|||||:|||||
Db 629 GlyAlaMetGlyThrLeuArgAsnSerIleGlnAspAsnThrAlaThrLysAsnGlyGln 648
QY 694 AATAGCTTATCACTATTAAATGAACAAAAGACT----- 726
Db :|||||:|||||
Db 649 AsnTyLeuAspAlaThrGluArgAsnLysThrAsnTyAsnAsnAlaValAspSerAla 668
QY 727 -----AATGCTGATGCTATTCTAATATAGTTTATTT 756
Db :|||||:|||||
Db 669 AsnGlyValIleAsnAlaThrSerAsnProAsnMetAspAlaAsnAlaIleAsnGlnIle 688
QY 757 AAAAAAGTGATCAAAATAATCAACAAAGTTTTGTAGGACCTTTTACAAACCGCTAATGTT 816
Db :|||||:|||||
Db 689 AlaThrGlnValThrSerThrLysAsnAlaLeuAspGlyThrHisAsnLeuThrGlnAla 708
QY 817 CAACCTTCAACACTACAGTTTTGTTGCTGCTGATGATGATCAACCCCGCTCAATATAAA 876
Db :|||||:|||||
Db 709 LysGlnThr-----AlaThrAsnAlaIleAspGlyAlaThrAsnLeuAsn 723
QY 877 TATGCAAGAAGCAGCGCTTTGGTGAATGCTGATGACCTTCAAGTAGAATTTTTCGAACACG 936
Db :|||||:|||||
Db 724 LysAlaGlnLysAspAlaLeuLysAlaGlnValThrSerAlaGlnArgValAlaAsnVal 743
QY 937 AATAGTATCACAGATGTTTCTGGATTATATAGTTTAGCT----- 975
Db :|||||:|||||
Db 744 ThrSerIleGlnGlnThrAlaAsnGlnLeuAsnGlnLeuGlnMetGlyGlnLeuGlnHisGly 763
QY 976 -----GGAAACAACACGAAGTACCAGTTAGTTTGTAGCAACTATGCTCCA 1020
Db :|||||:|||||
Db 764 IleAspAspGluAsnAlaThrLysGlnThrGlnLysTyArgAspAlaGluGlnSerLys 783
QY 1021 TCAACTGGTATTATTATTTCCCTTATAAGTTGGTTAAAGCTGATGCT----- 1071
Db :|||||:|||||
Db 784 LysThrAlaTyI-----AspGlnAlaValAlaAlaAlaLysAlaIleLeuAsn 799
QY 1072 -----AATACGTTGGATTACATAAATAAAT----- 1101

Db 800 LysGlnThrGlySerAsnSerAspLysAlaValAspArgAlaLeuGlnValThr 819
 QY 1102 -----AATGGAATGTTCAACAAGTTGAGTTTGCCACCTCAACTAGT 1143
 Db 820 SerThrLysAspAlaLeuAsnGlyAspAlaLysLeuAlaGluAlaLysAlaAlaArg 839
 QY 1144 GCAAAAT-----AATACTACAGTAAATCCAACTCCAGCAGTTGATGAG 1185
 Db 840 GlnAsnLeuGlyThrLeuAsnHisIleThrAsnAlaGlnArgThrAlaLeuGluGlyGln 859
 QY 1186 ATTAAGTGTGTAATAATCGTTTATCAGTTTAAAGATTGGCCAAACACAAATCGAATTA 1245
 Db 860 IleAsnGlnAlaThrThrVal-----AspGlyValAsnThrValLysThr 874
 QY 1246 AGTGTTCACAGCGGTGAAGGAATATGAAAT-----AAGATT 1281
 Db 875 AsnAlaAsnThrLeuAspGlyAlaMetAsnSerLeuGlnGlyAlaIleAsnAspLysAsp 894
 QY 1282 CGGCCAATGATGGCAACATTTATCTTAGCTCAATGAAAT 1323
 Db 895 AlaThrLeuArgAsnGlnAsnTyrLeuAspAlaAspGluSer 908

RESULT 10

fibrinogen-binding protein - *Staphylococcus aureus*

N:Alternate names: clumping factor

C:Species: *Staphylococcus aureus*

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999

C:Accession: S41539; S36630

Mol. Microbiol. 11, 237-248, 1994

A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*

A:Reference number: S41539; PMID:94224142; PMID:8170386

A:Accession: S41539

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 <MCD>

A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Alignment Scores:

Pred. No.: 0.00139 Length: 933

Score: 173.50 Matches: 111

Percent Similarity: 35.90% Conservative: 66

Best Local Similarity: 22.52% Mismatches: 185

Query Match: 7.46% Indels: 131

DB: 2 Gaps: 23

US-09-147-052-1 (1-1371) x S41539 (1-933)

QY 4 CACTATTTTAGCGGAATTCATATTTTCTCTTATAGTTATCTA----- 48
 Db 9 HisAlaIleArgLysLysSerIleGlyValAlaSerValLeuValGlyThrLeuIleGly 28
 QY 49 TATGGTACGAACTCATCT-----CCAGTACCCAAATGTGACATCAAGAA 96
 Db 29 PheGlyLeuSerSerLysGlyAlaAspAlaSerGluAsnSerValThrGlnSerAsp 48
 QY 97 GTTGTTCGAGCGCTCCAGTGTCTGAGGAAGAGTCTACGTCTTATCTTTGTCCTCCCA 156
 Db 49 SerAlaSerAsnGluSerLysSerAsnAspSerSerSerValSerAlaAlaProLysThr 68
 QY 157 GTGGTTCACCGCTGATCCGCTCTAGAAATTCGGCTGTATCTGTCTATTAATAAAGATCA 216
 Db 69 AspAspThrAsnVal-----SerAspThrLysThrSerSer 80
 QY 217 AACCCAAATAAGGCCAAACCAATTA-----GAAGCAGCGGAATGAGTAAACAGAT 270
 Db 81 AsnThrAsnAsnGlyGlnThrSerValAlaGlnAsnProAlaGlnGlnGlnThrGln 100
 QY 271 CTAATCAATGCTAAACGATGACATAGCTTCACTACAAGACTATGCCAAGATTGAGCT 330
 Db 101 SerSerSerThrAsnAlaThrThr----- 108

QY 331 AGTTTATCACTGCTTATAGTCAAGCTCAACACAGTTAACATAACCTTAATGCAACATTA 390
 Db 109 GluGlnThrProValThrGlyGluAlaThrThrThrThrThrThrThrThrThrThrThr 128
 QY 391 GAACAACATAAAATGGCTAAACCTAAATTTAGAAATCAGCCATCAACCAAGCTAATACCGAT 450
 Db 129 AlaThrThrGlnSerSerAsnThrAsnAlaGluGluLeuValAsnGln-----ThrSerAsn 147
 QY 451 AAAACGACTTTTATTAATGAACCCCAATTTAGTTGAAGCATACAAAGCACTAAACAC 510
 Db 148 GluThrThrPheAsnAsp-----ThrAsnThrValSerSerValAsnSer----- 162
 QY 511 ACTTTGAACAACGCTGCTACTTAACCTTGAAGGTTTGTCTCACTCACT----- 555
 Db 163 -----ProGlnAsnSerThrAsnAlaGluAsnValSerThrThrGlnAspThrSerThr 180
 QY 556 -----GCTTATAATCAAAATTCGCAATTAATTTAGTGGATCTATACAATAA--- 600
 Db 181 GluAlaThrProSerAsnAsnGluSerAlaProGlnSerThrAspAlaSerAsnLysAsp 200
 QY 601 -----GCTAGTAGTTTAACT 618
 Db 201 ValValAsnGlnAlaValAsnThrSerAlaProArgMetArgAlaPheSerLeuAlaAla 220
 QY 619 AAAACACATGATCCACTAAATGGGGAACGCTTTTAGATTCTTAATGAGATTACTACAGCT 678
 Db 221 ValAlaAlaAspAlaProAlaAlaGlyThrAspIle-----ThrAsnGlnLeuThrAsnVal 239
 QY 679 AATAAGCAATTAATAATAGTTTATCACTTAACTAATTAATGAACA---AAGACT 726
 Db 240 ThrValGlyIleAspSerGlyThrThrValTyrProHisGlnAlaGlyTyrValLysLeu 259
 QY 727 AAT---GCTGATGCTATCTATAGTTTATTAATA---AAGTGATTCAA 771
 Db 260 AsnTyrGlyPheSerValProAsnSerAlaValLysGlyAspThrPheLysIleThrVal 279
 QY 772 AATAATGACAAAGTTTGTAGGACTTTTACAAACGCTAATGCTTCAACCTTCAAACTAC 831
 Db 280 ProLysGluLeuAsnLeuAsnGlyValThrSerThrAlaLysValProPro----- 296
 QY 832 AGTTTGTGTTTGTAGTCTGATGTACACCCGCTCAATTAATATCAAGAGGACC 891
 Db 296 ----- 296
 QY 892 GTTTGGAATGGTATGAACTTCAAGTAGAATTTCTGCAACACGAAATAGTAGTATCAGAT 951
 Db 297 IleMetAlaGlyAspGln-----ValLeuAlaAsnGlyValLysAspSerAsp 312
 QY 952 GTTCTTGTGATTATAGTTTACTGGA---ACAAACACGAAGTACCAATTTAGTTTATGAC 1008
 Db 313 GlyAsnValIleTyrThrPheThrAspTyrValAsnThrLysAspValLysAlaThr 332
 QY 1009 AACTATGTCCTCACTCACTGTTTATTTATATTTCCCTTATAGTTGGTTAAACACGCTGAT 1068
 Db 333 LeuThrMetProAla-----TyrIle-----Asp 340
 QY 1069 GCTAATAAGTTGGATTACATAATAATAATGGAATGTTTCAACAAAGTTGAGTTT 1128
 Db 341 ProGluAsnVal-----LysLysThrGlyAsnVal----- 350
 QY 1129 GCCACTTCAACTAGTGCATAATTAATCACTACAGTAAATCACTCAACGCTGATGATGAGAT 1188
 Db 351 ThrLeuAlaThrGlyIleGlySerThrThrAlaAsnLysThrValLeuValAspTyrGlu 370
 QY 1189 AAGTTTGTCTAAATC-----GTTTATCATCAGTTTAAAGTTTGGC 1227
 Db 371 LysTyrGlyLysPheTyrAsnLeuSerIleLysGlyThrIleAspGlnIleAspLysThr 390
 QY 1228 CAACAACACATCGAATTAAGTTTCCCAAGGGTGAAGGAAATATGATAAAGTTGCGCCA 1287
 Db 391 AsnAsnThrTyrArgGlnThrIleTyrValAsnProSerGlyAspAsnValIleAlaPro 410

Db 1018 ValThrGluIleThrLysLeuGluThrPheAsnLysAspLysValLysPheArg 1037
::: ||||| ||||| :::
QY 664 GAGATTACTACAGCTAATAAGAATTAATAATAACGTATCAACTATTAATGAACAAG 723
||||| ||||| ||||| :::
Db 1038 GlueGluGlnThrArgLysAspAluPheIleasnThr-----AsnLysThrAsn 1056
||||| ||||| ||||| :::
QY 724 ACTAATGCTGATGCATTATCATATAGTTTAAAAAAAGTGATTCAAATAATGAACAA 783
||| ||| ::| ||| ||||| :::
Db 1057 ProAsnTyrSerThrLeuIleSerGluLeuThrSerLys-----ArgaspSerLysAsn 1074
||||| ||||| ||||| :::
QY 784 AGTTTTGTAGGAGCTTTTACAAAGCTAAATGTTCCAACTTCAAACTACAGTTTGTGGCT 843
||| ::| :||| :||| :||| :||| :|||
Db 1075 SerIleThrAsnSerSerAsnLysSerAspIleGluThrAlaAsnThrGluLeu----- 1092
||| ||| ::| ||| ||||| :::
QY 844 TTTAGTCTGATGTPAACACCCTCAATTATAATATCAAGAGGACCGTTTGGAAATCGT 903
||| ||| ::| ||| ||||| :::
Db 1092 ----- 1092
QY 904 GATGAACCTTCAAGTAGAATCTTCCAACACAGAAATAGTATCACAGATGTTCTTGGATT 963
::| ||||| ||| ||||| ||||| :::
Db 1093 -----LysGlnAlaLeuAlaLysAlaAsn-----ThrAspLysAlaGlnAla 1106
||| ||||| ||||| ||||| ||||| :::
QY 964 TATAGTTTGTAGCGAACAACACAGAGTACCAATTTAGTTTTCACAACTATGTGCCATCA 1033
:||| ||||| ||||| ||||| ||||| :::
Db 1107 AspAsnLeuAla---ArgSerThrLysGluGlnLeuAsnLysSerIleSerSerAlaAsn 1125
||| ||||| ||||| ||||| ||||| :::
QY 1024 ACTGGTTATTTATATTTCCCTTATATAGTTGGTTAAACACAGCTGATCTAATAACGTTGGA 1083
||| ||| ||||| ||||| ||||| ||||| :::
Db 1126 Thr-----LeuLeuAlaLysLeuThrAspLysAspAsnThrIle 1138
||| ||||| ||||| ||||| ||||| ||||| :::
QY 1084 TTACAATACAAATATAATATGGAATGTTCAACAAGTTGAGTTGCGACCTTCAACTAGT 1143
||| ||||| ||||| ||||| ||||| ||||| :::
Db 1139 GlnGlnAlaLysThrGlu-----LeuGluLysGluValGlnLysAlaAsnGlnAlaVal 1156
||| ||||| ||||| ||||| ||||| ||||| :::
QY 1144 GCAATTAATCTACAGCTAATFCAACCTCCACGAGCTTGATGAGATT-----AAAGTTGCT 1197
||| ||||| ||||| ||||| ||||| ||||| :::
Db 1157 AlaSerAsnAsnThrValSerMetGlnSerAlaLysSerLeuLeuAspThrLysValThr 1176
||| ||||| ||||| ||||| ||||| ||||| :::
QY 1198 AAAATCGTTTATCAGGTTTAAAGATTGGCCAAAACACAAATCGAATTAAGTTGTTCCAAAG 1257
||| ||||| ||||| ||||| ||||| ||||| :::
Db 1177 GluIleThrLysLysLeuGluThrPheAsnLysAspLys----- 1189
||| ||||| ||||| ||||| ||||| ||||| :::
QY 1258 GGTCGAAGGAATATGAATAAGTTGCGCAATGATGGCAACATT-----TATCTTAGC 1311
||| ||||| ||||| ||||| ||||| ||||| :::
Db 1190 --GluAlaLysPheAsnGluLeuLysLysThràrgGlyGlnIleGlnGluPheIleasn 1208
||| ||||| ||||| ||||| ||||| ||||| :::
QY 1312 TCAATGAAATATGCTGATAGATCCCCGGGTAC 1347
:||| ||||| ||||| ||||| ||||| ||||| :::
Db 1209 ThrAsnLysAsnAsn-----ProAsnTyr 1216
||| ||||| ||||| ||||| ||||| ||||| :::

RESULT 12
C89921
hypothetical protein ebhb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimura,
C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: C89921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <R>
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BAB42528.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhb

Alignment Scores:

```

Pred. No.: 0.00359 Length: 3890
Score: 166.50 Matches: 100
Percent Similarity: 37.20% Conservative: 89
Best Local Similarity: 19.69% Mismatches: 210
Query Match: 7.16% Indels: 109
DB: 2 Gaps: 18

US-09-147-052-1 (1-1371) x C89921 (1-3890)
QY 49 TATGGTACGAACATCATCTCCGAGTACC----- 75
Db 2427 HisSerValSerSerAsnProSerThrLeuThrAlaProAlaHisThrValAsnThr 2446
QY 76 -----CAAAATGTGACATCAAGAGAAAGTGTTCGAGCGTC 111
Db 2447 ThrGluLeuValLysAspTyrGlySerAsnValThrAlaAlaGluLeuAsnAlaVal 2466
QY 112 CAGTTGTCTGAGGAAGTGTCTACGTTTATCTTTGTCCTCCACCAGTGGTTCACACGGT 171
Db 2467 GlnValAlaAsnLysArgThr-----AlaThrIle 2476
QY 172 -----ATCCGCTCTAGAATTCGGCTGTATGCTATTACTATAAAAA 210
Db 2477 LysAsnGlyThrAlaMetProThrAsnLeuAlaGlyGlySerThrThrIleProVal 2496
QY 211 GATCAAAACCAATAATATGCGCAAC----- 237
Db 2497 ThrValThrTyrAsnAspGlySerThrGluGluValGlnGluSerIlePheThrLysAla 2516
QY 238 -----CAATTAGAACGCGCGAATGAGTTAACAGATCTAATCAATGCTAAAGCG 288
Db 2517 AspLysArgGluLeuIleThrAlaLysAsnHisLeuAspAspProValSerThrGluGly 2536
QY 289 ATGACATTACTCTCAGTACAGACTAT-----GCCAAGATTGAAGCTAGTTTA 336
Db 2537 LysLysProGlyThrIleThrGlnTyrAsnAsnAlaMetHisAsnAlaGlnGlnIle 2556
QY 337 TCATCTGCTTATAGTGAAGCTGAACAGTTAAACAATTAACCTTAATCAACATTAGACAA 396
Db 2557 AsnThrAlaLysThrGluAlaGlnGlnValIleAsnAsnGluArgAlaThrProGlnGln 2576
QY 397 CTAAATGCTTAACTAATTTAGAACTCAGCCCAACCAAGCTAATACGGATAAAGC 456
Db 2577 ValSerAspAlaLeuThrLysValArgAlaAlaGlnThrLysIleAspGlnAlaLysAla 2596
QY 457 ACTTTTGATTAATCAACACCAAT-----TTAGTTCAAGCATACAAACCACTAAACCC 510
Db 2597 LeuLeuGlnAsnLysGluAspAsnSerGlnLeuValThrSerLysAsnAsnLeuGlnSer 2616
QY 511 ACTTTAGACACAGTGTCTACTAACCTTGAAGTTTGTCTCATCACTGCTTATAATCAAAAT 570
Db 2617 SerValAsnGln-----ValProSerThrAla-----GlyMet 2627
QY 571 CGCAATAATTTAGTGGATCTATACAT-----AAAGCTAGTAGTTTAATAACT 618
Db 2628 ThrGlnGlnSerIleAspAsnTyrAsnAlaLysLysArgGluAlaGluThrGluIleThr 2647
QY 619 AAACACTAGATCCACTAATAGGGAACCGCTTTTAGATTCTTAATCAGATTACTACAGCT 678
Db 2648 AlaAlaGluArgValIleAspAsnGlyAsp--AlaThrAlaGlnGlnIleSerAspGlu 2666
QY 679 ATATAGATATTAATAATACGTTATCACTAATTAATGAACAAAGAACTAATGCTGATGCA 738
Db 2667 LysHisArgValAspAsnAlaLeuThrAlaLeuAsnGlnAlaLysHisAspLeuThrAla 2686
QY 739 TTATCTAATAGTTTATTAATAAAGTATTCAAAATATGAACAAGTTTGTAGGGACT 798
Db 2687 AspThrHisAlaLeuGluGlnAlaValGlnGlnLeuAsnArg-----ThrGlyThr 2703
QY 799 TTATCAAAAC-----GCTAATGTTCAACCTTCAACCTTCAACAGTTTGTGCTTTTACT 849
Db 2704 ThrThrGlyLysLysProAlaSerIleThrAlaTyrAsnAsnSerIleArgAlaLeuGln 2723

```

```

QY 850 GCTGATGTAACA-----CCGCTC-----AAT 870
Db 2724 SerAspLeuThrSerAlaLysAsnSerAlaAsnAlaIleLeuGlnLysProIleArgThr 2743
QY 871 TATAAATATCAAGAAGACCGTTTGAATGGTGATGAACCTTCAAGTACGAATTCITGCA 930
Db 2744 ValGlnGluValGlnSerAlaLeuThrAsnValAsnArgValAsnGluArgLeuThrGln 2763
QY 931 AACACGAATAGTATCACAGATGTTTCTGATGTTTATAGTTAGCTGGAACAAACAGAAAG 990
Db 2764 AlaIleAsnGlnLeuValProLeuAlaAspAsnSerAlaLeuArgThrAlaLysThrLys 2783
QY 991 -----TACCAATTTAGTTTATAGCAACTATGGTCCATCACTCACTGTTATTATAT 1038
Db 2784 LeuAspGluLeuLeuAsnLysSerValThrThrAspGlyMetThrGlnSerIleGln 2803
QY 1039 TTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTTAATACGTTGGATTACAATAACAATTA 1098
Db 2804 AlaTyrGluAsnAlaLysArgAlaGlyGlnThrGluThrThrAsnAlaGlnAsnValIle 2823
QY 1099 AATAATGGAATGTTCAACAAGTTGAGTTGGCTTCCACTCA----- 1137
Db 2824 AsnAsnGlyAspAlaThrAspGlnGlnIleAlaAlaGluLysThrLysValGluGluLys 2843
QY 1138 ---ACTAGTGCAATAATACTACAGCTAATCCAACTCCAGCAGTGTGATGAGATTAAAGTT 1194
Db 2844 TyrAsnSerLeuLysGlnAlaIleAlaGlyLeuThrProAspLeuAlaProLeuGlnThr 2863
QY 1195 GCTAAATCGTTTATCAGGTTTAAGATTGGCCAAACACACATCGAATTAAGTGTCCA 1254
Db 2864 AlaLysThrGlnLeu-----GlnAsnAspIleAsp-----GlnPro 2875
QY 1255 ACGGGTGAAGAAATATGAATAAGTTGCGCCCAATGATTGGCAACATTTATCTTACTGCTCA 1314
Db 2876 ThrSerThrThrGlyMetThrSerAlaSerValAlaAlaPheAsnAspLysLeuSerAla 2895
QY 1315 AATGAANAATAATGCTGATAAGATC 1338
Db 2896 AlaArgThrLysIleGlnGluIle 2903

RESULT 13
F82885
hypothetical protein UU482 [Imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: F82885
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4688 <GLA>
A:Cross-references: GB:AF222894; GB:AF222894; NID:96899476; PIDN:AAF30894.1; GSPDB:GN
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU482
A:Genetic code: SGC3

Alignment Scores:
Pred. No.: 0.00356 Length: 4688
Score: 166.50 Matches: 112
Percent Similarity: 35.78% Conservative: 78
Best Local Similarity: 21.09% Mismatches: 166
Query Match: 7.16% Indels: 175
DB: 2 Gaps: 25

US-09-147-052-1 (1-1371) x F82885 (1-4688)
QY 55 ACGAACACTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAG 114
Db 1387 SerAsnAsnAlaValSerLeuLysAsnAsnThrAsnIleGluIleThrLysArgIleLeu 1406

```

```
QY 115 TTGCTGAGGAGAGCTCTACGTTT-----TAT 141
Db 1407 ValAsnAsnGlnSerLeuIleSerAlaLysIleGluValAspAspIleAspAsnVal 1426
QY 142 CTTTGTCCCCACCAGTGGTTCACCGTGCCTGCTAGAAATCGGCTGCTGCTATT 201
Db 1427 LeuAsnThrAlaAspLysProAsnIleValTyrGlnLeuGlu----- 1440
QY 202 ACTAAAGAGTGAACCCCAAT-----AATGCCAAACCCCAATTAGAA 246
Db 1441 AsnSerAsnAspAlaAsnAsnAsnLeuLysLeuAlaThrAsnAlaGlnValValGlu 1460
QY 247 GCAGCGCAATGGAGTTAAACA--GATCTAATCAATGCTAAGCGATCACATTAGCTTCA 303
Db 1461 AsnAsnLysLysPheLeuLysPheAspLeuValSerLeuLysIleAsn----- 1476
QY 304 CTCAAGACTATGCCAAGATTGAAGCTAGTTTATCATCT----- 342
Db 1477 --GlnAsnTyrValIleLysGluIleSerPheAsnSerLysProThrAsnAlaTyrPhe 1495
QY 343 -----GCTTATAGTGAAGCTGAA----- 360
Db 1496 AsnPheThrAsnAsnLysThrAsnAsnIleValTyrSerTyrAspGluGlnAsnLysIle 1515
QY 361 ACAGTTACATAATACCTTAATGCAACA----- 387
Db 1516 SerLeuSerAsnAsnIleAsnProThrSerTyrLysProGlnThrIleAsnLysAsnAsp 1535
QY 388 -----TTAGAACAACCTAAATAATGGCTGCTAAACTAAT 417
Db 1536 LysSerValAsnIleAspValAspLeuGlnValAspLysGlnLeuLeuAlaAsnGlnTyr 1555
QY 418 TTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACT----- 459
Db 1556 LeuArgLeuLysLeuLysGlnLeuAsnAspAsnLysThrValTyrThrAspProIleLeu 1575
QY 460 TTTGATTAATGACACCCCAATTTAGTTGAACATCAACAAGCACATAAAACCACTTTAGAA 519
Db 1576 PheAsnAsn-----AsnAlaLysIleSerPheLys-----LeuSerAsnLeuIleHis 1591
QY 520 CAAGCTGTCTACTAACTTGAAGTTTGTGCATCACTGCTTATTAATCAATTCGCAATTAAT 579
Db 1592 AsnArgAlaTyrGluLeuGluGluLeu-----TyrTyrPheAspAspGlnAsnSer 1608
QY 580 TTAGTGGATCTATCAATAAAGCTAGTAGTTTAATACTAAACACACATAGATCCACTAAT 639
Db 1609 ValAsnAspMetThrAsnAsnGlnIleSerPheAsnSerLysIleHisLysPro----- 1626
QY 640 GGGGGAACGCTTTAGATTCTAATGAGATTACTACAGCTAATAGATAATTAATAATACG 699
Db 1627 -----LysIleGluPheGluProSerLeuThrThrIleAsnTyrAspThrAsnAla 1644
QY 700 TTATCAACTATTAAAT-----GAACAAAGACACTAATGCTGAT 735
Db 1645 IleLysThrValSerAlaHisAsnAlaGlnValHisPheLysLeuLysThrAsnAspGlu 1664
QY 736 GCATTATCTAATAGTTTATTATTAATAAAGTGTATCAAAATAATAGCAAAAGTTTGTAGG 795
Db 1665 AlaLeuGluAsnAspGlnIleValGluAlaVal----- 1675
QY 796 ACTTTTACAACGCTAATGTTTCAACCTTCAACCTTACAGTTTGTGCTTTTGTAGTCTGAT 855
Db 1675 ----- 1675
QY 856 GTAACACCCGCTAATTATAATATGCAAGAGGACCGTTTGGAAATGGTGAACCTTCA 915
Db 1676 PheAlaProThrAsnAsnLeuAsnAspGlnLysValVal-----Glu 1689
QY 916 AGTAGAATCTTGTCAACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCT 975
Db 1690 AlaLysLeuAsnAsnValThrAsnSerPheAsnGluGluGluLeuGluPheAsnLysSer 1709
```

```
QY 976 GGA-----ACAAACACGAGTACCATT-----AGTTTTAGCACTATGCTCCA 1020
Db 1710 GlyLeuArgGluGluThrThrTyrArgLeuIleLysValThrPheLysAsnLys----- 1727
QY 1021 TCAACTGTTTATATATTTCCCTTATTAAGTTGGTTAAAGCAGCTGATGCTTAATAACGTT 1080
Db 1728 -----ProAsnLysAlaTyrGluLeuLeuAsnLysAsnGlyVal 1740
QY 1081 GGATTACAATAAATAATGAAATGTTCAACAAGTTGAGTTTGCCACT----- 1134
Db 1741 IlePheGluTyrLys-----AsnGlySer---GlnAlaTyrGluPheThrThrGlnLys 1757
QY 1135 -----TCAACTAGTCAATAATAACTACACGCTAAT 1164
Db 1758 PheGluHisLysValIleAspValValSerSerThrSerThrAsnThrThrGlnGlnGlu 1777
QY 1165 CCAACTCCAGCAGTTGATGAGATT-----AAAGTTGCTAAAATCGTT 1206
Db 1778 IleThrValLysIleAspGlyIleGlnArgAlaTyrAsnAsnLysLysLeuGluLeuVal 1797
QY 1207 TTATCAGTTTAAGATTGGCCAAACACATCGAATTAAGTTTCCACGGGTGAAGGA 1266
Db 1798 TyrGluSerAsnIleLeuGlyAspProGluIleLysThrVal-----AspAsn 1814
QY 1267 AATATGAATAAGTTGCCCAATGATTCGCAACATTTATCTTAGCTCAATGAAATAAT 1326
Db 1815 AsnAsnAsn-----SerValHisLeuSerPheAspLysLysGlu 1827
QY 1327 GCTGATAGATCCCGGGTACCGTCGACCCGGT 1359
Db 1828 TyrAsnLeuValLeuAsnAsnLeuLysProGly 1838

RESULT 14
AG2422
hypoetical protein all14935 [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #test_change 09-Dec-2002
C:Accession: AG2422
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-661 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB76634.1; PID:gl7134073; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all14935

Alignment Scores:
Pred. No.: 0.00415 Length: 661
Score: 166.00 Matches: 106
Percent Similarity: 32.91% Conservative: 50
Best Local Similarity: 22.36% Mismatches: 216
Query Match: 7.14% Indels: 102
DB: 2 Gaps: 17

US-09-147-052-1 (1-1371) x AG2422 (1-661)
QY 1 ATGCACATATTTAGCGGGAATTCATATTTTTCCTTATA-----GTTATTCTATATGTT 54
Db 1 MetAsnLeuValThrGlnLysAlaIleLeuGlyLeuIleThrThrAlaAlaLeuThrGly 20
QY 55 ACGAACTCAVCT-----CCGAGTACCCAAAATGTCACATCAAGAGAACTT 99
Db 21 ThrLeuAlaSerIleLysProAlaThrAlaThrThrGlnThrThrThrSerAlaAsn--- 39
QY 100 GTTTTCGAGCTCCAGTTGCTCTGAGGAAGAGCTACGTTTATCTTTTGTCTCCCCACCAGTG 159
Db 100 GTTTTCGAGCTCCAGTTGCTCTGAGGAAGAGCTACGTTTATCTTTTGTCTCCCCACCAGTG 159
```

```
Db 40 ValLysProIleGlnValAlaIleLysGluAla-----ProGluAlaIle 54
QY 160 GGTTCACCGGTATCCGCTAGAAVTCGGCTGTATCTATTACTATAAAAGATGCAAA 219
Db 55 -----LysLys----- 56
QY 220 CCAATAATAGGCCAAACCAATTAGAACACGCGGAATGGAGTTAAACAGATCTAATC--- 276
Db 57 AlaGlnGluAlaAlaAspAlaLeuLysLysAlaGluAspLeuSerGlyIleIleArg 76
QY 277 -----AATGCTAAAGCGATGACATTAGCTTCACCTACAGACATAT 315
Db 77 SerThrAsnArgAlaLysSerAspAlaAsnSerArgLeuAsnArgAlaGluGlnAspPhe 96
QY 316 CCCAAGATTGAAGCTAGTTTATCATCTGCTTATATAGTGAAGCTGAACAGATTAAACAATAAC 375
Db 97 ThrGlnAlaAspAlaGlyPheGlnThrAlaGlnThrAlaLeuAsnThrAlaThrAsnAsn 116
QY 376 CTTAATGCAACATTAGAACAACTAAATAATGGCTAAACTAATTTAGATCAGCCATCAAC 435
Db 117 ArgAsnAsnAlaGlnAsnAlaLeuAsnThrAlaThrAsnAsnArgAsnAsnAlaGlnAsn 136
QY 436 CAAGCTAATACGGATAAAGACGTTTTCATAATGAACACCAACCAATTTAGTTGAAGCATAC 495
Db 137 AlaLeuAsnThrAlaThrAsnAsnArgAspAsn-----AlaGln 149
QY 496 AAAGCACTAAACCACTTTAGAACACACGCTGCTACTAACCTTGAAGTTTGTGCATCAACT 555
Db 150 AsnAlaLeuAsnThrAlaThrAsnAsnArg---AsnAsnAlaGlnAsnAlaLeuAsnThr 168
QY 556 GCTTATATCAAAATTCGAATAATTTAGTGGATCTATACAAATAAGCTAGTAGTTTAATA 615
Db 169 AlaThrAsnAsn---ArgAsnAsnAlaGlnAsnAlaLeuAsnThrAlaThrAsnAsnArg 187
QY 616 ACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATCTTAATGAGATTACTACA 675
Db 188 AsnAsnAlaGlnAsnAlaLeu-----AsnThrAlaThrAsn 199
QY 676 GCTAATAAGAAATATTATATAGCTTATCACTAATTAATGAACAAAGACTAATGCTGAT 735
Db 200 AsnArgAsnAsnAlaGlnAsnAlaLeuAsnThrAlaThrAsnAsnArgAsnAsnAlaGln 219
QY 736 GCATTATCTAATAGTTTATTAATAAAGTGATTCAA-----AATAATGAACAAAGT 786
Db 220 AlaGluValAspThrAlaThrArgAsnLeuAlaGlnAlaArgArgGlyAsnSerGlnLys 239
QY 787 TTTGTAGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTT 846
Db 240 AlaIleGlnAsnAlaIleAsnAlaLeuAsnGlnAlaAsnThrArgLeuAsnThrAlaAsn 259
QY 847 AGTCTGATGTACACCCCGTCAATTAATAATATGCAAGAGGACCGCTTGGATGCTGAT 906
Db 260 ThrAlaLeuAsnThrAlaThrAsnAsnPheAsnThrAlaAsnThrAlaPheAsn----- 277
QY 907 GAACCTTCAAGTAGAATCTTGCACACAGATAGTATCAGAGATGTTCTTGGATTAT 966
Db 278 -----ThrAlaAsnAsnAsnPheSerThr----- 285
QY 967 AGTTAGCTGGAACAAACAGAGTACCAATTTAGTTTTCAGCAACTATGCTCCATCAACT 1026
Db 286 -----AlaAsnThrAlaPheAsnThrAlaThrAsnAsnPheSerThrAla----- 300
QY 1027 GGTATTATTATTCCTTTATAAGTTTGGTTAAAGCAGCTGATGCTTAATACCTTGGATT 1086
Db 301 -----AsnThrAlaLeu 304
QY 1087 CAATACAAATTAATAATGGAATGTTCAACAAAGTTGAGTTTGGCCACTTCAACTAGTACA 1146
Db 305 AsnThrAlaThrAsnAsnPheAsnThrAlaIleAlaGluLeuAspGlnAlaAsnThrArg 324
QY 1147 AATAATACTACGCTAATCAACTCCAGCAGTTGATGATGATGATGATGATGATGATGATGAT 1206
Db 325 LeuAsnThrAlaArgAsn-----AspPheAsnThrAlaAsnSerAsn 338
```

```
QY 1207 TTATCAGGTTTAAGATTGGCCAAACACACATCGAATTAAGTGTTCACACGGGTGAAGGA 1266
Db 339 PheSerArgThrGlyAsnGluLeuAsnThrAlaThrAsnAsnPheAsnThrAlaAsnAsn 358
QY 1267 AATATGAATAAAGTTGGCCCAATGATTGGCAACATTTATCTAGCTCAAAATGAAATAAT 1326
Db 359 ThrPheAsnThrAlaThrAsnPheAsnAsn-----AlaSerSerArgArgAsnThr 376
QY 1327 GCTGATGAAGATCCCGGGTACCGTCCGACCGGTACATTTTAA 1368
Db 377 AlaGluGlnAlaArgAsnGlnValArgGluGluThrArgLeu 390
RESULT 15
D90011
FmtB protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cul, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D90011
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2481 <KUR>
A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BA843253.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fmtB(mrp)
Alignment Scores:
Pred. No.: 0.00422 Length: 2481
Score: 165.50 Matches: 104
Percent Similarity: 35.86% Conservative: 66
Best Local Similarity: 21.94% Mismatches: 173
Query Match: 7.12% Indels: 131
DB: 2 Gaps: 22
US-09-147-052-1 (1-1371) x D90011 (1-2481)
QY 73 ACCCAAAATGTGACATCAAGAAAGTTGTTTCGACGGTCCAGTTGTTCTGAGGAAGAGTCT 132
Db 1150 ThrGlnAspIleThrAlaGluGluIleAlaAla----- 1161
QY 133 ACGTTTATCTTTGTCCTCCACCCACCGTGGTTCAACCGGTGATCCGCTAGATTCGGCTGT 192
Db 1162 -----AsnAlaAspValAspAsnAlaValThrGlnAlaAsnSerAsnIle 1176
QY 193 ATGTCTATTACTAAAAA-----GATGCAAAACCCCAATAATATGCGCAACCACTTA 273
Db 1197 AspGlnValThrProThrValAsnLysLysAlaThrAlaArgAsnGluIleThrAlaIle 1216
QY 274 ATCAATGCTAAAGCGATGACATTAGCTTCACCTACAGACTATGCCAAGATTGAAGCTAGT 333
Db 1217 LeuAsnAsnLysLeuGlnGluIleGlnAlaThrProAspAlaThrAspGluGluLysGln 1236
QY 334 TTATCATCTGCTTATAGTGAAGCTGAAACAGTCAACATTAACCTAAT----- 381
Db 1237 AlaAlaAspAla-----GluAlaAsnThrGluAsnGlyLysAlaAsnGlnAlaIleSer 1254
QY 382 -----GCAACATTAGAACAACTAAATGCTCAAACTAAATTTAGATCAGCCATCAAC 435
Db 1255 AlaAlaThrThrAsnAlaGlnValAspGluAlaLysAlaAsnAlaGluAlaAlaIleAsn 1274
QY 436 CAAGCTAATACGGATAAAACGACTTTTGATTAATGAACACCACTTATAGTTGAAGCATAC 495
```



```
Db      1275 ---AlaValThrProLysValValLysGlnAlaAlaLysAspGluIleAsnGlnLeu 1293
QY      496 AAGCAGCTAAACACCTTTA-----GAACAACGTGCTACTTAACCTTGAAGGTTTG 546
Db      1294 GlnAlaThrGlnThrAsnValIleAsnAsnAspGlnAsnAlaThrAsnGluGluLysGlu 1313
QY      547 TCA-----TCAACGCTTATATCAATTCGAATAATTTA-----582
Db      1314 AlaAlaIleGlnGlnLeuAlaThrAlaValThrAspAlaLysAsnAsnIleThrAlaAla 1333
QY      583 -----GTGCACTATACAATAAAGCTAGTCTTAATAAATAAACAACACTA 627
Db      1334 ThrAspAspAsnGlyValAspThrAlaLysAspAlaGlyLysAsnSerIleGlnSerThr 1353
QY      628 GATCCACTAAATGGGGGACGCTTTTAGATTCTAAT-----GAG 666
Db      1354 GlnPro-----AlaThrAlaValLysSerAsnAlaLysAsnGluValAspGlnAla 1370
QY      667 ATTACTACAGCTAATAAGATATTAATAATACGTTATCACTATTAATGAACAAAGACT 726
Db      1371 ValThrThrGlnAsnGlnAlaIleAspAsnThrThrGlyAlaThrThrGluGluLysAsn 1390
QY      727 AATGCTGATGCATTATCTAATAGTTTATTAAAGAGTTCAAATAATGAACAAAGT 786
Db      1391 AlaAlaLysAspLeu-----ValLeuLysAlaLysGluLysAla 1403
QY      787 TTTGTAGGAGCTTTTACAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTT 846
Db      1404 TyrGlnAspIleLeu-----AsnAlaGlnThrThrAsn-----1414
QY      847 AGTGTGATGTAACACCGCTCAATTAATAATATCAAGAAGACCGTTTGGAAATGGTAT 906
Db      1415 -----AspValThrGlnIle-----1419
QY      907 GAACCTTCAAGTAGAATCTTGGCAACACGAATAGTATCACAGATGTTTCTTGGATTAT 966
Db      1420 -----LysAspGlnAlaValAlaAspIleGlnGlyIleThrAlaAspThrIleLys 1437
QY      967 AGTTTAGCT-----GGACAACACGAGTAGTACCAATTTAGTTTGTAGCAACTAT 1014
Db      1438 AspValAlaLysAspGluLeuAlaThrLysAlaAsnGluGlnLysAlaLeu-----1454
QY      1015 GGTCCATCACTGGTTATTTATATTTCCCTTATAGTTGGTTAAGCAGCTGATGCTAAT 1074
Db      1455 -----IleAlaGlnThrAlaAspAlaThr 1462
QY      1075 -----AAGCTTGGATTACAATAACAATAAATAATGGAATGTT 1113
Db      1463 ThrGluGluLysGluGlnAlaAsnGlnGlnValAspAlaGlnLeuThrGlnGlyAsn---1481
QY      1114 CAACAAGTTGAGTTTGGCACTTCAACTAGTGCATAATATACTACAGCT---AATCCAAC 1170
Db      1482 GlnAsnIleGluAsnAlaGlnSerIleAspAspValAsnThrAlaLysAspAsnAlaIle 1501
QY      1171 CCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTTTAAGATTTGGCCAA 1230
Db      1502 GlnAlaIleAspProIleGlnAlaSerThrAspValLysThrAsnAlaArg-----1518
QY      1231 AACACAATCGAATTAAGTTGTTCCAAAGGTCGAAGGAATATGAATAAAGTTGCCCAATG 1290
Db      1519 -----AlaGluLeuLeuThrGluMet-----GlnAsnLysIleThrGluIle 1532
QY      1291 ATTGGCAACATTTATCTTAGCTCAATGAATAATATGCTGAT 1332
Db      1533 LeuAsnAsnAsnGluThrThrAsnGluLysGlyAsnAsp 1546
```

Search completed: October 8, 2003, 18:17:13
Job time : 82.2146 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:56:57 ; Search time 15.327 Seconds
(without alignments)
5245.316 Million cell updates/sec

Title: US-09-147-052-1
Perfect score: 2326
Sequence: 1 atgcactattttaggcgaa.....gaccgggtacattttataa 1371

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 137520 seqs, 29319821 residues

Total number of hits satisfying chosen parameters: 275040

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09147052/runat_08102003_154343_29933/app_query.fasta_1.5980
-DB=Pending Patents_AA_New -OPWT=fastan -SUFFIX=rapn -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -THR.WAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052 -CGN_1_12=runat_08102003_154343_29933 -NGPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending_Patents_AA_New.*
2: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	156	6.7	807	5	US-09-820-843B-108
2	142.5	6.1	770	5	US-09-897-516A-4453
3	138	5.9	1010	6	US-10-415-657-10
4	134	5.8	861	5	US-09-820-843B-109
5	133	5.7	523	6	US-10-425-114A-58966
6	124.5	5.4	1225	5	US-09-581-286A-485
7	124.5	5.4	1226	5	US-09-581-286A-484
8	124.5	5.4	1251	5	US-09-581-286A-351
9	124	5.3	615	6	US-10-650-274-166
10	123	5.3	930	5	US-09-200-650E-3
11	123	5.3	3241	6	US-10-647-057-1

12	122	5.2	386	1	PCT-US03-20322-229	Sequence 229, App
13	122	5.2	1531	1	PCT-US02-37235-44	Sequence 44, Appl
14	121	5.2	945	1	PCT-US03-20460-14	Sequence 14, Appl
15	120	5.2	1315	5	US-09-200-650E-5	Sequence 5, Appl
16	119.5	5.1	1386	6	US-10-472-078-8	Sequence 8, Appl
17	119.5	5.1	1449	6	US-10-472-078-2	Sequence 2, Appl
18	119	5.1	1033	5	US-09-820-843B-75	Sequence 75, Appl
19	118	5.1	1881	1	PCT-US03-27401-316	Sequence 316, Appl
20	117.5	5.1	477	6	US-10-425-114A-71726	Sequence 71726, A
21	117.5	5.1	1222	6	US-10-650-274-142	Sequence 142, App
22	117	5.0	524	6	US-10-425-114A-70497	Sequence 70497, A
23	117	5.0	918	5	US-09-200-650E-1	Sequence 1, Appl
24	116.5	5.0	990	6	US-10-047-676B-7	Sequence 7, Appl
25	116	5.0	741	6	US-10-425-114A-64315	Sequence 64315, A
26	115.5	5.0	1032	1	PCT-US03-20460-4	Sequence 4, Appl
27	115	4.9	1097	6	US-10-650-274-111	Sequence 111, App
28	114	4.9	402	1	PCT-US02-34769-6	Sequence 6, Appl
29	114	4.9	403	1	PCT-US02-34769-14	Sequence 14, Appl
30	114	4.9	647	6	US-10-425-114A-5165	Sequence 5165, A
31	114	4.9	725	5	US-09-897-516A-5773	Sequence 5773, Ap
32	114	4.9	1222	6	US-10-650-274-103	Sequence 103, App
33	113.5	4.9	660	1	PCT-US02-35624-139	Sequence 139, App
34	112	4.8	880	6	US-10-425-114A-70095	Sequence 70095, A
35	112	4.8	1419	1	PCT-US02-35624-173	Sequence 173, App
36	112	4.8	2538	1	PCT-US03-20082-49	Sequence 49, Appl
37	111.5	4.8	402	6	US-10-425-114A-38936	Sequence 38936, A
38	111.5	4.8	1554	1	PCT-US03-06962-38	Sequence 38, Appl
39	110.5	4.8	441	6	US-10-425-114A-39801	Sequence 39801, A
40	109.5	4.7	904	1	PCT-US03-11231-18	Sequence 18, Appl
41	109.5	4.7	1153	5	US-09-897-516A-7395	Sequence 7395, Ap
42	109.5	4.7	2599	5	US-09-674-546A-1669	Sequence 1669, Ap
43	109	4.7	1518	1	PCT-US03-06962-40	Sequence 40, Appl
44	108.5	4.7	742	7	US-60-487-610-2540	Sequence 2540, Ap
45	108.5	4.7	1274	5	US-09-581-286A-466	Sequence 466, App

ALIGNMENTS

RESULT 1
US-09-820-843B-108
; Sequence 108, Application US/09820843B
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843B
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patent version 3.2
; SEQ ID NO 108
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: hypothetical protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: g1|3845292
US-09-820-843B-108

Alignment Scores:
Pred. No.: 0.000278
Score: 156.00
Percent Similarity: 39.41%
Best Local Similarity: 23.78%
Query Match: 6.71%
DB: 5
Length: 807
Matches: 73
Conservative: 48
Mismatches: 124
Indels: 62
Gaps: 14

US-09-147-052-1 (1-1371) x US-09-820-843B-108 (1-807)

Oy 364 GTTAACTAACCTTAATGCAACATTAGACAACTAAATAACT----- 414


```
QY 739 TTA---TCTAATAGTCTTTTAAATAAAGTGAATCAAAATATACAAAGTTTGTAGG 795
Db 629 ProAspLysAsnAlaPheValIleGlyAsnIle-----GlyLeuGly 641
QY 796 ACTTTTACAAACCGTAAATGTTTCAACCTTCAAACTACAGTTTGTCTTTAGTGTCTGAT 855
Db 642 AspLeuIleGlySerLysIleGluSerGln-----LeuIleGlyGlnAsp 656
QY 856 GTACACCCCGTCAATATTAAATATGCAAGAAGGACCGTTTGGATGGTGATGAACCTTCA 915
Db 657 AlaThr----- 658
QY 916 AGTAGAATCTTGCACACCAAGTAGATATACAGATGTTCTTGGATTATAGTTAGCT 975
Db 659 -----IleIleAsnLeuGlyIleThrGlnIleSerGlyVal---AlaIleAla 674
QY 976 GGAACAAACAGCAAGTACCAATAGTTTAACTAGCACTATGTCCTCAACTGCTTATTTA 1035
Db 675 GlyThrProIleLys---GlnGluAsnThrSerIleValGlyGlyValThr-Tyr-TyrThr 693
QY 1036 TATTTCCTTTATAGTTGGTTAAAGCAGCTGATGCTAATACGTTGGATTACATACAAA 1095
Db 694 AsnTyr---TyrLys-----IleArgLeuProValSer 703
QY 1096 TTAATAATGGAATGTTCAACAAGTTGAGTTTGCACCTTCAACTAGTGCAAAATAATACT 1155
Db 704 LeuProAsnGly-----IleIleSerCysHisAlaSerIleAlaCysAsnAsnPhe 720
QY 1156 ACAGCTAATCCAACTCCAGCAGTGTGATGAGATTAAAGTTGCTAAAGTCTTTTATCAGGT 1215
Db 721 AspSerGlnLeuProSerHisLeuAlaAspValArgThrGlnArgSerAsnSerAspGly 740
QY 1216 TTAGATTGGCCAAACACAAATCGAATTAGTTGTTCCAAACGGGTGAA 1263
Db 741 ValGlyLeuSerGlnAspThrLeuThrIleSerValThrProGlu 756

RESULT 3
US-10-415-657-10
; Sequence 10, Application US/10415657
; GENERAL INFORMATION:
; APPLICANT: Genix Bioscience GmbH
; TITLE OF INVENTION: Eukaryotic cell division genes and their use in diagnosis and tre
; TITLE OF INVENTION: of proliferative diseases
; FILE REFERENCE: CE61773US
; CURRENT FILING DATE: 2003-09-08
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1010
; TYPE: PRN
; ORGANISM: C. elegans
US-10-415-657-10

Alignment Scores:
Pred. No.: 0.00588 Length: 1010
Score: 138.00 Matches: 107
Percent Similarity: 40.83% Conservative: 100
Best Local Similarity: 21.10% Mismatches: 164
Query Match: 5.93% Indels: 136
DB: 6 Gaps: 28

US-09-147-052-1 (1-1371) x US-10-415-657-10 (1-1010)

QY 61 TCATCTCCGAGTACCCAAAATGTCATCATCAGAGAAGTGTTCGAGGTCGAGTGTGCT 120
Db 363 AsnSerProLysHisValSerLysSerMetAspLeuGluLysThrIleGlu---Ala 381
QY 121 GAGGAAGAGTCTACGTTTATCTTGTCCCCACAGCGGGTTCACCGTGATCCGCTCA 180
Db 382 AlaAspLysSerThrLysTyr-----ProSerIleAlaAspGluValGluAspLeu 398
```

```
QY 181 GAATTCGGCTCTATGCTTATTTACTTAAAAAAGAT-----GCAAAACCCAAATAAT 228
Db 399 AspMetAsp---MetAspIleThrGluGlnGlnProCysGluAlaGlyAsnGlnAsn 417
QY 229 GGCCAAACCCCAATTAGAGCAGCG---CGAATGGAGTTAACAGATCTCTAATCAATGCTTAA 285
Db 418 AspGlyLeuGlnLeuGlnLysGluAspLeuMetAspIleSerValIleArgAspSerPro 437
QY 286 GCGATG-----ACATTAGCTTCTACACAACTATGCAAGATGGAAGCTAGTTTATCA 339
Db 438 AlaValAsnAspThrMetAlaValPheGlnSerProAlaArgValLysIleGlyAlaAsn 457
QY 340 TCTGCTTATAGTGAAGCTGAA----- 360
Db 458 AsnSerIleIleAspSerGlnLysSerIleValPheGlyAspGluMetSerIleAspGlu 477
QY 361 ----- 372
Db 478 ThrGlnAsnAspGlyThrLeuThrLeuProLysSerAsnValGluValThrThrAsn 497
QY 373 AACCTTATGCAACATTAGAACAACTTAAATGGCTAAACCTAATTTAGAATCAGCCATC 432
Db 498 AspValTyrThrSerLeuGluArgGlnGluAsnAlaSerGluAsnValSerMetIle 517
QY 433 AACCAAGCTAATAGC-----GATAAACGACTTTTGTGATAATGAACACCCAAT 480
Db 518 AsnGluSerSerValHisSerGluIleAspLysLysSerPhe-----Met 532
QY 481 TTAGTTGAAGCATCAAAAGCA---CTAAAAACCACTTTAGAACACGCTACTAACCTT 537
Db 533 LeuIleGluGluArgAlaPheMetHisSerMetIleAspValAlaGlnLysLeu 552
QY 538 GAA-----GGTTGTCTCACTCACTTATATCAATCAAAATTCGCAATAATTTAGTGATCTA 591
Db 553 GluAspAspGlySerSerLysThrPro-----ValIleLeu 564
QY 592 TACAATAAGCTAGTGTCTTAACTATAA----- 621
Db 565 AlaSerGlnSerAlaSerLeuAlaThrLysGluProSerAlaLeuHisAsnSerSerAla 584
QY 622 ACCTAGAT-----CCACTAATGGGGAACGCTTTAGATTCTTAATGAGATTACT 672
Db 585 ThrLeuAsnAsnSerMetGluLeuAspAsnThrLeuLeuLysThrMetGlnIleThr 604
QY 673 ACAGCTAATAAGAATATTAAATACGTTATCACTATTATTAATGAACAA-----AAG 723
Db 605 ThrCys-----GluAspIleSerMetValHisGluSerIleAlaValGlu 619
QY 724 ACTAATGCTGATGCAATTATCTAATAGTTTATTAAAAAGCTGATTCAAAATAATGAACAA 783
Db 620 LeuAsnSerAsnLysGluGlnGluPheGlyAspGluThrLeuGlnLysAsnAspThr 639
QY 784 AGTTTGTAGG-----ACTTTT-----ACA 804
Db 640 SerAsnThrGlyAlaAsnPheThrPheGlnGlyHisAsnGluThrSerGlnIleMetAsn 659
QY 805 AACGCTAATGTTCAACCTTCAACCTACAGTTTGTGCT-----TTTAGTGTGCTGATTAACA 861
Db 660 AsnValAspSerGluAlaValAsnThrSerLysIleSerThrTyrSerAlaPheAsnLeu 679
QY 862 CCCGTCAATTATAATATGCA-----AGAGGACCGTTTGAATGGTGGATGAACCTTCA 915
Db 680 SerIleAsnGlnSerIleSerLysArgArgSerLeuLeuAsnSerAlaArgGluSer 699
QY 916 AGTAGAATCTTGCACACCAAGTAGATATCAGATGTTCTTGGATTATAGTTTACT 975
Db 700 ProArgValAlaLeuGluAsn-----IleMetSerMetAsn 713
QY 976 GGAACAAAC-----ACGAAGTACCAATTTAGTTTTAGCACTATGGTCCATCA 1023
Db 714 GlyGlnThrMetGluAlaLeuThrGluTyrArg-----GlnAsnLysThrMetGln 730
```

QY	415	AAATTTAGAA-----TCAGCCATCAACCAAGCTAAT-----	444
Db	464	ValAlaGluLeuLeuMetLysGlnIleSerMetIleLysGluArgAsnLysGlyLeuAsp	483
QY	445	-----ACGGATAAAGCAGCTTTT-----GATAATGAACACCCCAATATTAGTTGAA	489
Db	484	ValLeuGluLysLysAsnThrPheGlyPheLeuAspAsnAsnTyrlGlnAsnTy-	501
QY	490	GCATACAAAGCACATAAAACCCTTTAGAACACAGTGCTACTAACCTTGAAGGTTTGTC	549
Db	502	-----GlySerAsnAsnAsnSer	507
QY	550	TCAACTGCTTATAATCAAAATTCCCAATAATTTAGTGGATCTATACAAATAAAGCTAGTAGT	609
Db	508	SerLeuGluLysAsnAsnMetLysGluAsn-----AspIleTyrSerLysGluAlaSer	525
QY	610	TTA-----ATAACTAAACACATAGATCCACTAAATGGG-----GGA	645
Db	526	LysArgIleMetAspIlePheArgThrLeuAsnSer--AsnGlyLeuValSerGlnGlu	544
QY	646	ACGCTTTTAGATCTCAATGAGATTACTACAGCTAATAAGAATATTATAATACGTTATCA	705
Db	545	SerLeuLeuValAsnGlnSerValLeuAsnAsnAsnAsnAsnTyrlAsnAsnTyrlAsnSer	564
QY	706	ACTATTATGAACAAAGAGCTAATGCTGATGCATCTATTCTAATAGTTTTTATTAATAAAGTG	765
Db	565	AsnAsnAsnArgAsnLys--AsnGlnAsnAsnAsnAsnAsnAsnAsnMetAsn	583
QY	766	ATTCAAAATATGAACAAGATTTTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCA	825
Db	584	AsnMetAsnAsnSerAsnAsnAsnIleAsn-----AsnAsnAsnAsnTyrlAsnAsn	601
QY	826	AACTACAGTTTGTGCTTTTAGTCTGATGTAACACCCGCTCAATTATAAA-----	876
Db	602	AsnHisLysTyr-----HisSerMetAspAsnValThrTyrLysLysIlePhe	617
QY	877	-----TATGCAAGAGGACCGCTTTTGGATGGTGATGACCTTCAAGTAGAATTCCT	927
Db	618	IleAsnAsnTyrSerAsnAsn-----AspGlyAsnAsnAsnSerAsn-----	631
QY	928	GCAACACGAATAGTATCACAGATGTTTCYTGATTTATAGTTTACGTGGACAAACACAG	987
Db	632	--AsnSerAsnSerAsnAsnValGluHisTyrTyr-----MetAsnAsn	646
QY	988	AAGTACCAATTT-----AGTTTAGCAACTATGCTCCATCAACTGTTTATTATATTC	1041
Db	647	LysLysAsnPheLysAsnLysIleAsnAsnTyr-----	657
QY	1042	CCTTATAAGTTGGTTAAGCAGCTGATCTAATACGTT-----GGATTACAATPAC	1093
Db	658	---HisAsnLeu-----ProAspAsnLysAsnAsnMetMetAsnAsnAsnThrTyrAsn	674
QY	1093	AAATTAATATATGAAATGTTTCAACAAGTTCCAG---TTTGCCACTTCAACTAGTGC	1149
Db	675	AsnIleAsnLysAsnAsnLeuSerAsnMetGluAsnPheProProSerLeuSerPheAsn	694
QY	1150	AATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAATCGTTT	1209
Db	695	AsnSerAspIleAsnLysAsnAsnAlaGlnGlyAsnIleAsnIleThrProIleIleAsn	714
QY	1210	TCAGGTTTAAGATTGGC-----CAAAACACAATCGAA-----	1242
Db	715	SerIleLeuArgLeuAspAsnGluValAspAsnValHisAsnAsnSerIleSerGluAsn	734
QY	1243	-----TTAAGTTGTCCA	1254
Db	735	IleGlnAsnAlaLysValSerAsnValLeuAspSerLeuLysLysAlaSer	754
QY	1255	ACGGGTGAAGAAATATGATAAAGTTGCGCCCAATGATTGGCACATTTATCTAGTCA	1314
Db	755	LysSerGlnGlyAsnAsnAsnTyrlAsnIleProLys-----AsnPheAsnAsnAsn	772

```

208 AAGATGCAACCCAAATAATGGCCAACCCTATTACAAGCAGCGCATGGAGTTAACA 267
    |||::: ||| ||||| |:::
390 LysAsnAsnAsnAsnAsnGlnYAsnAsnIleLeasn-----Asn 403
    |||   |||   |||   |||
268 GATCTAATCAATGCTAAAGCGTGACACTTAGTTCCTACTACAACACTATGCGCAAGATTGAA 327
    :::   |||   :::   :::: |||
404 AsnSerAsnAsnLysAsnAsnIlePheGlnGlyAsnSerArgAsnTyrGluAsnValMet 423
    :::   |||   :::   :::: |||
328 GCTAGT-----TTATCATCTCTTATAGTGAAGCT--- 357
    :::   ::::   :||:::   :||::: |||
424 TyrAsnIleAsnAsnAsnAsnAsnIleSerAsnAsnLysAsnGluAlaSer 443
    |||   |||   |||:::   |||::: |||
358 ---GAAACACTTAACCACTTATGCGCAACTTAGAACACTAAAATGGCTAAACT 414
    |||   |||   |||:::   |||::: |||
444 PheAsnThrAspAsnIleAsnThrAsnSerGlyArgGluGluLysIleSerAsnThr 463

```

Alignment Scores:		
pred. No.:	0.058	Length: 1225
Score:	124.50	Matches: 98

Percent Similarity: 32.89%
 Best Local Similarity: 21.63%
 Query Match: 5.35%
 DB: 5

Conservative: 51
 Mismatches: 155
 Indels: 149
 Gaps: 22

US-09-147-052-1 (1-1371) x US-09-581-286A-485 (1-1225)

```

QY 337 TCATCTGCTTATAGTGAAGCTGAACAGCTTAAACATAAC-----CTT 378
Db 358 AsnAlaAlaIleAsnGluAlaGlnPheAlaLysAsnSerGlyTyrThrIleHisThrIle 377
QY 379 AATGCAACATTAGAACAACTAAATGCTAAACTAATTTAGATAGATCAGCCATCAACAA 438
Db 378 GlyTyrAspLeuGlyAspPheAlaIleAlaAsnSerLeuLysLeuThrAla----- 395
QY 439 GCTAATACGGATAAAGACCTTTTGTATAATGAACACCCA-----AATTAGTTGAAGCATAC 495
Db 396 -----ThrAspGluAsnHisPhePheThrAlaThrProAlaAsnLeuAlaAlaPhe 413
QY 496 AAGCAGCTAAAAACACCTTTAGAACACCGTCTACTAACCTTGAAGTTGTGTCATCAACT 555
Db 414 AspAsnIleAlaGlnThrIleAsnIleGlyIleGlnArgGlyGluValThrAspPheVal 433
QY 556 GCTTATATCAAAATCGCAATAATTTAGTGGATCTATACAATAAGCTAGTGTATTAATA 615
Db 434 AlaProGlyPheIleValLysAsnLeuThrGln-----SerGlyAspVal 448
QY 616 ACTAAACACATAGATCCACTAAATGGG----- 642
Db 449 ThrHisLeuLeuAsnValSerAsnGlyThrValHisTyrAspValSerThrLysLysLeu 468
QY 643 -----GGAACGCTTTAGATTTCTAATGAGATTACTACAGCT----- 678
Db 469 ThrTrpThrThrGlyThrIleLeuSerSerSerGluAlaThrIleThrTyrArgIleTyr 488
QY 679 -----AATAGAATATTAAT---AATACGTTATCAACTAATTAAT 714
Db 489 AlaAspLeuAspTyrIleGlnAsnAsnAspIleProValAsnThrThrSerAlaIleGly 508
QY 715 GAACAAAGACTAATGCTGATCATATCTAATAGTTTATTAAAAAGTGATCAAAAT 774
Db 509 ProAspLeuGlyGlyPheAspThr-----Asn 517
QY 775 AATGAACAAAGTTTGTGAGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGT 834
Db 518 ThrGluAlaLysLeu-----ThrTyrThrAsnSerAsnGlyGluProAsnGlnGln--- 534
QY 835 TTTGTTGCTTTTGTAGTCTGATGTAACACCGCTCAATTATTAATATGCAAGAGGACCGTT 894
Db 535 ---LeuIlePheProArgProThrValLysLeuGlyTyrGlyValIleLysArgHisTyr 553
QY 895 TGGAAATGGTGAACCTTCAAGTAGAATTTCTGCAACACAGCAATAGTATCACAGATGTT 954
Db 554 ValLeuValAsnLysAspGlyGlnProIleGlnAlaAsnGlyThrValValSerSerLeu 573
QY 955 TCT----- 957
Db 574 SerGluAlaHisValLeuGlnSerGlnAspPhePheLeuProSerGlyGlyGlyHisIle 593
QY 958 -----TGGATT-----TATAGTTAGCT 975
Db 594 ValProLysTrpIleLysLeuAspLysThrThrGluAlaLeuGlnTyrTyrSerValPro 613
QY 976 GGACAAACACAG-----AAGTACCAATTT----- 999
Db 614 ProThrAsnThrValIleThrThrAlaAspGlyLysArgTyrArgPheValGluValPro 633
QY 1000 ---AGTTTACCACTATGCT-----CCATCAACTGGTTATTATTA 1035
Db 634 GlySerThrProAsnProGlyGlnIleGlyIleSerTrpLysLysProAlaGlyAsnAla 653
QY 1036 TATTTCCCTTATAAGTTGGTTAAA----- 1059

```

```

Db 654 TyrPheAlaTyrLysLeuLeuAsnTyrTrpMetGlyGlyThrThrAspGlnGlnSerGlu 673
QY 1060 ---GCACCTGATGCTAATAACGTTGGATTACATAACATAATTAATATGGAATGTTCAA 1116
Db 674 TrpAspValThrSerAsnTrpThrGlyAlaGlnValProLeuThrGly-----Glu 690
QY 1117 CAAGTTGAGTTTGCACCTTCAACTAGTCGCAAAATAACTACAGCTAATCAACTCCAGCA 1176
Db 691 AspValGluPheAlaThrThrGluAsnPheGly-----SerProAla 704
QY 1177 GTTGATGAGATTAAAGTT-----GCTAAATCGTTTATCAGGTTTAAAGATT 1224
Db 705 ValAlaAspLeuHisValProThrThrAsnProLysIleIle-----GlyAsnLeuIle 722
QY 1225 GGCCAAACACATCGAATTAAGTTTCCACCGGT-----GAAGGAAAT 1269
Db 723 AsnAsnSerAspLysAspLeuValValThrThrSerSerGlnLeuThrIleAsnGlyVal 742
QY 1270 ATGAATAAAGTTCGCCCAATGATTGCAACATTTATCTAGCTCAATGAAATAATGCT 1329
Db 743 ValGluAspAsnAsnProAsnValGlyThrIleValValLysSerSerLysAspAsnPro 762
QY 1330 GATAAGATC-----CCGGGTACCGTCGA 1353
Db 763 ThrGlyThrLeuLeuPheAlaAsnProGlyTyrAsnGln 775

```

RESULT 7.

```

US-09-581-286A-484
; Sequence 484, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTH, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 484
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (983)..(983)
; OTHER INFORMATION: xaa can be any naturally occurring amino acid

```


; ORGANISM: Porphyromonas gingivalis
 ; FEATURE:
 ; NAME/KEY: MOD_RES
 ; LOCATION: (1008)..(1008)
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
 ; US-09-581-286A-351

Alignment Scores:
 Pred. No.: 0.0581 Length: 1251
 Score: 124.50 Matches: 98
 Percent Similarity: 32.89% Conservative: 51
 Best Local Similarity: 21.63% Mismatches: 155
 Query Match: 5.35% Indels: 149
 DB: 22 Gaps: 22

US-09-147-052-1 (1-1371) x US-09-581-286A-351 (1-1251)

```

QY 337 TCATCTGCTTATAGTGAAGCTGAACACAGTTAAACAATAAC-----CTT 378
Db 384 AsnAlaAlaIleAsnGluAlaGlnPheAlaLysAsnSerGlyTyrThrIleHisThrIle 403
QY 379 AATCCACATAGAACAACTAAATAATGGCTAAACATAATTTAGAACAGCCATCAACCAA 438
Db 404 GlyTyrAspLeuGlyAspPheAlaLeuAlaAsnSerLeuLysLeuThrAla----- 421
QY 439 GCTAATACGGATAAACGACTTTTGATAATGAACACCCA---AATTTAGTTGAACATAC 495
Db 422 -----ThrAspGluAsnHisPhePheThrAlaThrProAlaAsnLeuAlaAlaPhe 439
QY 496 AAGCATAAAACCCACTTTAGAACACGCTGCTACTAAACCTTGAAGCTTTGTGCATCAACT 555
Db 440 AspAsnIleAlaGlnThrIleAsnIleGlyLeuGlnArgGlyGluValThrAspPheVal 459
QY 556 GCTTATAATCAAAATCCCAATAATTTAGTGGATCTATACATAAAGCTAGTACTTAATA 615
Db 460 AlaProGlyPheIleValLysAsnLeuThrGln-----SerGlyAspVal 474
QY 616 ACTAAACACTAGATCCACTAAATGGG----- 642
Db 475 ThrHisLeuLeuAsnValSerAsnGlyThrValHisTyrAspValSerThrLysLysLeu 494
QY 643 -----GGAACGCTTTAGATTCCTAATGAGATTAATCACTACACT----- 678
Db 495 ThrTrpThrThrGlyThrIleLeuSerSerSerGluAlaThrIleThrTyrArgIleTyr 514
QY 679 -----AATAAGATAATTAAT---AATAGCTTATCAACTAATAAT 714
Db 515 AlaAspLeuAspTyrIleGlnAsnAsnAspIleProValAsnThrThrSerAlaIleGly 534
QY 715 GAACAAAGACTATGCTGATGATTAATTAATAGTTTATTAAAAAGTGATTCAAAAT 774
Db 535 ProAspLeuGlyGlyPheAspThr-----Asn 543
QY 775 AATGAACAAAGTTTGTAGGACTTTTACAAACGCTAATGTTCAACCTTCAACACACAGT 834
Db 544 ThrGluAlaLysLeu-----ThrTyrThrAsnSerAsnGlyGluProAsnGlnGln--- 560
QY 835 TTGTGTGCTTTTAGTCTGATGCTAATACACCCGCTAATTAATAATATGCAAGAGGACCGTT 894
Db 561 ---LeuIlePheProArgProThrValLysLeuGlyTyrGlyValIleLysArgHisTyr 579
QY 895 TGGATGGTGATGACCTTCAAGTGAATCTTGGCAACAGCAATAGTATCACAGATGTT 954
Db 580 ValLeuValAsnLysAspGlyGlnProIleGlnAlaAsnGlyThrValValSerSerLeu 599
QY 955 TCT----- 957
Db 600 SerGluAlaHisValLeuGlnSerGlnAspPhePheLeuProSerGlyGlyHisIle 619
QY 958 -----TGGATT-----TATAGTTAGCT 975
Db 620 ValProLysTrpIleLysLeuAspLysThrThrGluAlaLeuGlnTyrTyrSerValPro 639

```

```

QY 976 GGAACAACACG-----AAGTACCAATTT----- 999
Db 640 ProThrAsnThrValIleThrThrAlaAspGlyLysArgTyrArgPheValGluValPro 659
QY 1000 ---AGTTTATGACCAACTATGGT-----CCATCAACTGCTTATTATA 1035
Db 660 GlySerThrProAsnProGlyGlnIleGlyLeSerTriplyLysProAlaGlyAsnAla 679
QY 1036 TATTTCCTTATAAGTTGGTTAA----- 1059
Db 680 TyrPheAlaTyrLysLeuLeuAsnTyrTrpMetGlyGlyThrThrAspGlnGlnSerGlu 699
QY 1060 ---GCAGCTGATGCTAATAACGTTGGATTACATAACAAATTAATAATGAATGTTCAA 1116
Db 700 TipAspValThrSerAsnTrpThrGlyAlaGlnValProLeuThrGly-----Glu 716
QY 1117 CAAGTTGAGTTGCCACTTCACTAGTGCATAATAATACAGTAATCAACCTCCAGCA 1176
Db 717 AspValGluPheAlaThrThrGluAsnPheGly-----SerProAla 730
QY 1177 GTTGATGAGATTAAGCTT-----GCTAAATCGTTTATCAGGTTTAGATTT 1224
Db 731 ValAlaAspLeuHisValProThrThrAsnProLysIleIle-----GlyAsnLeuIle 748
QY 1225 GGCACAAACACATCGAATTAAGTGTCCACAGGGT-----GAAGGAAAT 1269
Db 749 AsnAsnSerAspLysAspLeuValValThrThrSerSerGlnLeuThrIleAsnGlyVal 768
QY 1270 ATGAATAAAGTTGCCCAATGATGTCACACATTTATCTAGCTCAATGAAATAATGCT 1329
Db 769 ValGluAspAsnAsnProAsnValGlyThrIleValLysSerSerLysAspAsnPro 788
QY 1330 GATAAGATC-----CCGGGTACCGTCGA 1353
Db 789 ThrGlyThrLeuLeuPheAlaAsnProGlyTyrAsnGln 801

```

RESULT 9

US-10-650-274-166
 ; Sequence 166, Application US/10650274
 ; GENERAL INFORMATION:
 ; APPLICANT: GLENN, MATTHEW
 ; APPLICANT: HAVUKKALA, ILKKA J
 ; APPLICANT: LUBBERS, MARK WILLIAM
 ; APPLICANT: DEKKER, JAMES
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,
 ; TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING
 ; FILE REFERENCE: 11000.1073
 ; CURRENT APPLICATION NUMBER: US/10/650,274
 ; NUMBER OF SEQ ID NOS: 187
 ; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
 ; SEQ ID NO 166
 ; LENGTH: 615
 ; TYPE: PRT
 ; ORGANISM: LACTOBACILLUS RHAMNOSUS
 US-10-650-274-166

Alignment Scores:
 Pred. No.: 0.0615 Length: 615
 Score: 124.00 Matches: 97
 Percent Similarity: 33.73% Conservative: 46
 Best Local Similarity: 22.88% Mismatches: 159
 Query Match: 5.33% Indels: 122
 DB: 6 Gaps: 20

US-09-147-052-1 (1-1371) x US-10-650-274-166 (1-615)

```

QY 55 ACGAACATCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTGAGCGTCCAG 114
Db 46 SerAsnSerSerSerThrThrAlaAsn-----LysSerValLys 58
QY 115 TTGCTGAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACACGAGTGGGTTCAACCGTATC 174

```

```
Db 59 Phe-----ProValSerTyrAsnAsnThrAlaLys 68
QY 175 GGTCTAGAAATCGGCTGTATGCTATTAATAAAGATGCAAAACAAATAATGGCCAA 234
Db 69 AlaIleLysGlyGlyAsnValAsnValAlaValValAsnAspSerProPheLysGly--- 87
QY 235 ACCCAATTAGACGACGCGAATGGAGTTACAGATCTAATCAATGCTTAAGCGATGACA 294
Db 88 -----ValPheAsnGluGluLeuTyrThr 95
QY 295 TTAGTCTCATTACAGACTATGCCAAGATTGAAGTAGTTTATCATCTGCTTATAGTAA 354
Db 96 AspAlaTyrAspAsnAspTyrMetSerProAlaAlaGluSerLeuPheAlaTyrAsnSer 115
QY 355 GCTGAACAGCTTAACAATAACCTTAATGCACATAGAACCACTAAATGGCTAAACT 414
Db 116 ThrPheLysPheAsnAsnAsnGlyAlaAlaThrIleLysGlnAspAsnSerAlaLysThr 135
QY 415 AATTAGATCAGCCATCAACCAAGCTAATACGGATAAAGCACTTTGTGTAATGAACAC 474
Db 136 -----IleThrValThrIleLysProAsnValLysTyrSerAspGlyGln 150
QY 475 CCA-----AATTAGTTGAAGCATACAAA-----GCACATAAAACCACTTTA 516
Db 151 ProValValAlaArgAspLeuValTyrAlaTyrGluIleMetAlaAsnLysAlaThrLys 170
QY 517 GAACACAGCTGCTACT-----AACCTTGAAGGTTTGTCTCACTCAACTGCTTATAAT 564
Db 171 SerGlnArgTyrThrGluSerLeuGlnAsnIleGluGlyLeu-----ThrGluTyrHis 188
QY 565 CAATTCCCAATAATTAGTGGATCATACATAAAGCTAGTAGTTTAATTAACATAACA 624
Db 189 Asp-----GlyLysAlaAspThrIleSerGlyLeuThr 199
QY 625 CTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAG 684
Db 200 MetProLysGlyAspAsnGly-----AsnThrMetValIleHisPheLys 214
QY 685 AATATAATAATACCTTATCACTAATTAATGAACAAAAGACTAATGCTGATGCAATATCT 744
Db 215 GlnMetLysProSerPheAsnThrSerGlyAsnGlyTyrPheLeuGluSerAlaAlaPro 234
QY 745 AATAGTTTATTAATAAAGTGNATCAAAATAATGAACAAAGT-----786
Db 235 TyrHisTyrLeuHisAspValAlaPheAspLysLeuGluSerSerAspLysValArgLys 254
QY 787 -----TTGTAGGACTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTT 837
Db 255 GlnProLeuPhePheGlyProTyrLysIleSerLysValValAlaGlyGln-----271
QY 838 GTTGCTTTTACGTGATGTAAACACCGCTCAATTAATAATATGCAAGAAGGACCGTTTGG 897
Db 272 -----SerValGluTyrThrProAsnGlnTyr-----TyrTrp 282
QY 898 ATGTGTATGAACCTTCAAGTAGAATT---CTTGCAACACGAATAGTATCACAGATGT 954
Db 283 LysGlyLysProSerLeuSerLysIleThrPheGluAsnValSerSerAlaSerIleThr 302
QY 955 TCTTGGATTATAGTTTACGTGGAACAAACAGAAAGTACCAATTT-----999
Db 303 Ser-----AlaLeuLysAsnHisLysTyrAspIleValTyrGlyMetPro 317
QY 1000 -----AGTTTACCAACTATGGT---CCATCA 1023
Db 318 SerAspSerTyrSerAspTrpLysAsnIleSerGlyTyrThrAsnLeuGlyHisGlnSer 337
QY 1024 ACTGGTTATTTATATTTCCCTATTAAGTTGGTTAAGCAGCTGAT-----GCTAATAC 1077
Db 338 LeuAlaTyrAsnTyrLeuGlyPheLysLeuGlyLysTyrPheAspLysLysSerGluAsn 357
QY 1078 GTT-----GGATTACAAATACAAA 1095
```

```
Db 358 ValTyrAspProAsnSerLysMetAlaAsnLysSerLeuArgGlnAlaMetGlyTyrAla 377
QY 1096 TTAATAATAATGAAATGTTCAACAAGTTGAGTTTGGCACTTCAACTAGTGCAAAATAACT 1155
Db 378 LeuAsnAsnAspGlnVal---AlaAlaLysPheTyrAsnGlyThrArgSerArgAlaThr 396
QY 1156 ACAGCTAATCCA 1167
Db 397 ThrLeuIlePro 400

RESULT 10
US-09-200-650E-3
; Sequence 3, Application US/09200650E
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eibhinn, Deirdre M.
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aure
; FILE REFERENCE: P062830US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-3

Alignment Scores:
Pred. No.: 0.074 Length: 930
Score: 123.00 Matches: 90
Percent Similarity: 34.80% Conservative: 60
Best Local Similarity: 20.88% Mismatches: 167
Query Match: 5.29% Indels: 114
DB: 5 Gaps: 18

US-09-147-052-1 (1-1371) x US-09-200-650E-3 (1-930)
QY 157 GTGGTTCAACCGTATCGCTCTAGAAATTCGGCTGTATGCTATTACTAAAAAGATGCA 216
Db 35 ValGlyThrThrLeuIle-----PheGlyLeuSerGlyHisGluAlaAla 51
QY 217 AACCCAAATAATGGCCAAACCCCAATAGAACGCGCGAATGGAGTTAACA-----267
Db 52 GluHisThrAsnGly-----GluLeuAsnGlnSerLysAsnGluThrThrAlaProSer 69
QY 268 -----GATCTAATCAATGCTAAACCGATGACATTAGTTCA 303
Db 70 GluAsnLysThrThrLysValAspSerArgGlnLeuLysAspAsnThrGlnThrAla 89
QY 304 CTACAAGACTATCCCAAGATTGAAGCTAGTTTATCATCTGCTATAGTGAA---GCTGAA 360
Db 90 ThrAlaAspGlnProLysValThrMetSerAspSerAlaThrValLysGluThrSerSer 109
QY 361 ACAGTTACAAATAACCTTAATGCACANTTAGAACAACTAAAAATGGCTAAACTAATTTA 420
Db 110 AsnMetGlnSerProGlnAsnAlaThrAlaAsnGln-----SerThrThrLys 125
QY 421 GAATCAGCCATCAACCAAGCTAATACGGATAAAGCACTTTTGATAATGAA-----CAC 474
Db 126 ThrSerAsnValThrThrAsnAspLysSerSerThrThrThrThrThrThrThrThr 145
QY 475 CCAATATTTAGTTCAAGCATACAAAGCA-----CTAAAAACCACTTTAGAACAA 522
Db 146 SerAsnLeuThrGlnAlaLysAspValSerThrThrThrThrThrThrThrThrThr 165
```

```
QY 523 CGTGCTACTAACCTTGAAGTTTGTCACTCACTGCTTATATCAATCAATTCGCAATTAATTA 582
Db 166 ArgThrLeuAsnArgMetAlaValAsnThrValAlaAlaProGlnGlnGlyThrAsnVal 185
QY 583 GTGGATCTATACAATAAAGCTAGTAGTTTAACTAAACACTAGATCCCAATAATGGG 642
Db 186 AsnAspLysValHisPheSer
QY 643 GGAACGCTTTAGATTCTAATGAGATT-----ACTACAGCTAATAAGAAATATTAAT 696
Db 195 AspIleAlaIleAspLysGlyHisValAsnGlnThrThrGlyLysThrGluPheThrPala 214
QY 697 AGTTTATCACTATTAATGACAAAGACTAAT-----
Db 215 ThrSerAspValLeuLysLeuLysAlaAsnTyrThrIleAspSerValLysGlu 234
QY 729 -----AGTTT-----
Db 235 GlyAspThrPheThrPheLysTyrGlyGlnTyrPheArgProGlySerValArgLeuPro 254
QY 730 -----GCTGATGCATTAATCTAATAGTTTATTAATAAAGATGATCA 771
Db 255 SerGlnThrGlnAsnLeuTyrAsnAlaGlnGlyAsnIleAlaLysGlyIleTyrAsp 274
QY 772 AATATGAACAAGATTTGTAGGACTTTTACAAAGCTAATGTTCAACCTTCAACTAC 831
Db 275 SerThrThrAsnThrThrTyrThrPheThrAsnTyrValAspGlnTyrThrAsnVal 294
QY 832 -----AGTTT-----
Db 295 ArgGlySerPheGluGlnValAlaPheAlaLysArgLysAsnAlaThrThrAspLysThr 314
QY 841 GCTTTTAGTGTGATGTAACACCGCTCAATTAATATATCAAGAGCGTTTGGAA 900
Db 315 AlaTyrLysMetGluValThrLeuGlyAsnAspThrTyrSerGluGluIleIleValAsp 334
QY 901 ---GGTGATGAACCTTCAAGTAGAATCTTGCACACAGCAATAGTATCAAGAT----- 951
Db 335 TyrGlyAsnLysLysAlaGlnProLeuLysSerSerThrAsnTyrIleAsnAsnGluAsp 354
QY 952 -----GTTTCTGGATTATAGTTTACCTGGACACAAACAGAACTACCAA 996
Db 355 LeuSerArgAsnMetThrAlaTyrValAsnGlnProLysAsnThrTyrThrLysGlnThr 374
QY 997 TTTAGTTTACCACTACTGCTCCATCACTGGTTATTTATATTTCCCT----- 1044
Db 375 PheValThrAsn-----LeuThrGlyTyrLysPheAsnProAsnAlaLysAsn 390
QY 1045 TATAAGTTGTTAAGCAGCTGATGCTAATAACGTTGGATTACAAATTAATAAT 1104
Db 391 PheLysIleTyrGluValThrAspGlnAsnGlnPheValAspSerPheThrProAspThr 410
QY 1105 GGAATCTTCAACAAGTT-----GAGTTTCCCACTCACTAGTCAATAATATACATA 1158
Db 411 SerLysLeuLysAspValThrAspGlnPheAspValIleTyrSerAsnAspAsnLysThr 430
QY 1159 GCTAATCCAACTCCAGCAGCTGATGAGATTA 1191
Db 431 Ala-----ThrValAspLeuMetLys 437

RESULT 11
US-10-647-057-1
; Sequence 1, Application US/10647057
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGAPPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
```

```
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3241
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-1

Alignment Scores:
Pred. No.: 0.0776 Length: 3241
Score: 123.00 Matches: 85
Percent Similarity: 37.73% Conservative: 61
Best Local Similarity: 21.96% Mismatches: 161
Query Match: 5.23% Indels: 80
DB: Gaps: 16

US-09-147-052-1 (1-1371) x US-10-647-057-1 (1-3241)
QY 187 GCCTGTATGCTTACTATAAAGATGCAACCAATATATGCGCAACCAATTAGAA 246
Db 1431 GlyAlaValSerValAsnLysIleAsp-----AsnAspValGluAlaSerValAsp 1447
QY 247 GCAGCGGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTA 306
Db 1448 LysSerSerIleGluGlyAlaAsnGluIleAsnValIleAlaLysAspValLysGlySer 1467
QY 307 CAAGCATATGCCAAGATTGAAGCTAGTTTA-----TCATCTGCTTATAGTGAA 354
Db 1468 SerAspLeuAlaLysGluTyrGlnAlaLeuAsnGlyLysAspLysLysTyrLeuGlu 1487
QY 355 GCTGAACAGCTTAACATAACCTTAATGCA-----ACATTAGAACAACATAAATGGCT 408
Db 1488 AspArgGlyIleAsnThrThrGlyAsnGlyTyrThrLysGluGlnLeuGluLysAla 1507
QY 409 AAAACTAAT-----TTAGAATCAGCCATCAACCAAGCTATATACGGATATA 453
Db 1508 LysLysLysGluGlyAlaValIleValAsnAlaLeuSerValAlaGlyThrAspLys 1527
QY 454 ACCACTTTTGAT-----AATGAACACCCCAATTTAGTTGAACA----- 492
Db 1528 SerAlaGlyGlyValAlaIleAlaValAsnThrValLysAsnLysPheLysAlaGluLeu 1547
QY 493 -----TACAAAGCACTAATAACCACTTTAGAACACAGCTGCTACTACCTGAA 540
Db 1548 SerGlySerAsnLysGluAlaGlyLysIleHisAlaLysHisValAsnValGlu 1567
QY 541 GGTTCCTCATCACTGCTTATATCAATTCGAATTCGAATTTAGTGGATCTATACAATAA 600
Db 1568 AlaLysSerSerThrValValValAsnAlaLeuSerGlyLeuAlaIleSerLysAspAla 1587
QY 601 GCTAGTAGTTTATAATACTAAACACTAGATCCCAATAATGGGGAACGCTT-----TTA 654
Db 1588 PheSerGlyMetGlySerGlyAlaTyrGlnAspLeuSerAsnAspThrIleAlaLysVal 1607
QY 655 GATTCTTAATGAGATTACT-----ACAGCTATAGAATAATTAATACGTTATCAACT 708
Db 1608 AspLysGlyArgIleSerAlaAspSerLeuAsnValAsnAlaAsnSerIleLeuGly 1627
QY 709 ATTAATGAACAAAGACTAATGCTGATGATTAAT-----AGTTTATT 756
Db 1628 ValAsnValAlaGlyThrIleAlaGlySerLeuSerThrAlaValGlyAlaAlaPheAla 1647
QY 757 AAAAAGTAGTTCAAAATTAATGAACAAAGT-----TTT 789
Db 1648 AsnAsnThrLeuHisAsnLysThrSerAlaLeuIleThrGlyThrLysValAsnProPhe 1667
QY 790 GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGATGTTTGTCTTTAGT 849
Db 1668 SerGlyLysAsnThrLysValAsnValGlnAlaLeuAsnAspSerHisIleThrAsnVal 1687
QY 850 GCTGATGTAACACCGCTCAATTTATAATATGCA-----AGAAGGACCGCTTTGG 897
Db 897
```

Db 1688 SerAlaGlyGlyAlaAlaSerIleGlyGlnAlaGlyIleGlyGlyMetValSerValAsn 1707
Qy 998 AATGGTGATGACCTTCAAGTAGAATCTTGGCAACACGAAAT----- 939
Db 1708 ArgGlySerAspGluThrGluAlaLeuValSerAspGluPheGluGlyValSerSer 1727
Qy 940 -----AGTATCACAGATGTTCTGGATTATATAGTTAGCTGGAACAACACGAG 990
Db 1728 PheAsnValAlaAspAlaLysAspGlnLysThrIleAsnThrIleAlaGlyAsn----- 1744
Qy 991 TACCAATTTAGTTTACCAACTATGGTCCCATCAACTGTTATTTATTTATTTCCCTTATAAG 1050
Db 1745 -----AlaAsnGlyGlyLysAlaAlaGly----- 1752
Qy 1051 TTGTTTAAAGCAGCTGATGCTAAT---AACGTTGGATTACAA----- 1089
Db 1753 ---ValGlyAlaThrValAlaHisThrAsnIleGlyLysGlnSerValIleAlaIleVal 1771
Qy 1090 -----TACAAATTAATAATGGAATGTTCAACAAGTTGAGTTGGCACTTCAACTAGT 1143
Db 1772 LysAsnSerLysIleThrThrAlaAsnAspGlnAspArgLysAsnIleAsnValThrAla 1791
Qy 1144 GCAATTAATACACGCTAAT 1164
Db 1792 LysAspTyrThrMetThrAsn 1798

RESULT 12

PCT-US03-20322-229
; Sequence 229, Application PC/TUS0320322
; GENERAL INFORMATION:
; APPLICANT: Dana-Farber Cancer Institute, Inc., et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; CURRENT APPLICATION NUMBER: PCT/US03/20322
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 229
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Bacillus cereus
PCT-US03-20322-229

Alignment Scores:
Pred. No.: 0.0847 Length: 386
Score: 122.00 Matches: 81
Percent Similarity: 37.28% Conservative: 64
Best Local Similarity: 20.82% Mismatches: 145
Query Match: 5.25% Indels: 99
DB: 1 Gaps: 14

US-09-147-052-1 (1-1371) x PCT-US03-20322-229 (1-386)

Qy 31 TTCCTTATAGTATCTATATAGTGACGAACATCATCTCCGAGTACCCAAATGTGACATCA 90
Db 19 PheIleProValSerAlaTyrAlaLysGluGlyGlnThrGluValLysThrValTyrAla 38
Qy 91 AGAGAAGTTGTTTCGCGGTCAGTGTCTGTCGAGGAAGAGTACGTTTATCTTGTGCC 150
Db 39 GlnAsnValIleAlaProAsnThrLeuSerAsnSerIleArgMetLeuGlySerGlnSer 58
Qy 151 CCACGAGTGGTTCACCGTGATCGGTCTAGAAATTCGGCTGTATGCTTACTATAAAAAA 210
Db 59 ProLeuIleGlnAla-----TyrGlyLeuIleIleLeuGlnGlnPro 72
Qy 211 GATGCAAAACCCAAAT-----AATGCCAAACCCCAATTAGAAGACGCG 252
Db 73 AspIleLysValAsnAlaMetSerSerLeuThrAsnHisGlnLysPheAlaLysAlaAsn 92
Qy 253 CGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAGAC 312

Db 93 ValArgGluTrpIleAspGluTyrAsnProLysLeuIleAspLeuAsn-----GlnGlu 110
Qy 313 TATGCCAAGATGAAGCTAGTTTATCATCTGTTAGTGAAGCTGAACACAGATTAAACAT 372
Db 111 MetMetArgTyrSerThrArgPheAsnSerTyrTyrSerLysLeuTyrGluLeuAlaGly 130
Qy 373 AACCTTATGCAACATTAGAA----- 393
Db 131 AsnValAsnGluAspGlnGlnAlaLysAlaAspPheMetSerAlaTyrGlyLysLeuGln 150
Qy 394 ---CAACTAAATGGCTAAACAACTAATTTAGAAATCAGCCATCAACCAAGCTAATACGGAT 450
Db 151 LeuGlnValGlnSerIleGlnGluSerMetGlnAspLeuGlnLeuAsnArgPhe 170
Qy 451 AAAACGAGCTTTTGTATATGAACACCCCAATTTA---GTTGAAGCATACAAAGACCACTAAA 507
Db 171 LysThrValLeuAspLysAspSerAsnAsnLeuSerIleLysAlaAspGluAlaIleLys 190
Qy 508 ACC-----ACTTTGAACAACAGCTGTACTAACCTTGA 540
Db 191 ThrLeuGlnGlySerSerGlyAspIleValLysLeuArgGluAspIleLysArgIleGln 210
Qy 541 GGTTCATCACTGCTTATATCAATTCGCAATAATTTAGTGGATCTATACATATAA 600
Db 211 Gly-----GluIleGlnAlaGluLeuThrThrIleLeuAsnArg 223
Qy 601 GCTAGTAGTTTAATA----- 615
Db 224 ProGlnGluIleIleLysGlySerIleAsnIleGlyLysGlnValPheThrIleThrAsn 243
Qy 616 -----ACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAAT 663
Db 244 GlnThrAlaGlnThrLysThrIleAspPheValSerIleGlyThrLeu-----SerAsn 261
Qy 664 GAGATTACTACAGCTAATAAGAATAATTAATACGTTATCAACTATTAAAGCAACAAAG 723
Db 262 GluIleValAsnAlaAlaAspSerGlnThrArgGluAlaAlaLeuArgIleGlnGlnLys 281
Qy 724 -----ACTAATGCTGTGATCATATCT 744
Db 282 GlnLysGluLeuLeuProLeuIleGlnLysLeuSerGlnThrGluAlaGluAlaThrGln 301
Qy 745 AATAGTTTATTAATAAAGTAGTTCATAAATAATGAACAAAGTTTGTAGGAGCTTTTACA 804
Db 302 IleThrPhe-----ValGluAspGlnValAsnSerPheThr-GluLeuIleAs 317
Qy 805 AACGCTATGTTCAACCTTCAACTACAGATTTTGTGCTTTTGTAGCTGTAGTGAACACCC 864
Db 317 pArgGlnIleThrThrLeuGluThrLeuLeuThrAspTrpLysValLeu-----AsnAs 335
Qy 865 GTCAATTATA-----AATATGCAAGAGGACCGTTTGGAAATGGTGATGAACCT 912
Db 335 nasnMetIleGlnIleGlnLysAsnValGluGluGlyThrTyrThrAspSerSerLeuLe 355
Qy 913 TCAGTAGAATCTTGGCAACACGAAATAGTATCACAGATGTTCTTGGATTATAGTTTA 972
Db 355 uGlnLysHisPheAsnGlnIleLysLysValSerAsp----- 367
Qy 973 GCTGGAACAACACCAAGTACCAAT 997
Db 368 ---GluMetAsnLysGlnThrAsn 374

RESULT 13

PCT-US02-37235-44
; Sequence 44, Application PC/TUS0237235
; GENERAL INFORMATION:
; APPLICANT: Handfield, Martin
; APPLICANT: Hillman, Jeffrey
; APPLICANT: Proqules-Fox, Ann
; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigen
; TITLE OF INVENTION: In the Diagnosis, Treatment, and Monitoring of Periodontal DI
; FILE REFERENCE: MBH01-662A

```

; CURRENT APPLICATION NUMBER: PCT-US02/37235
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Actinobacillus actinomycetemcomitans
PCT-US02-37235-44

Alignment Scores:
Pred. No.: 0.0893      Length: 1531
Score: 122.00         Matches: 102
Percent Similarity: 34.62%      Conservative: 60
Best Local Similarity: 21.79%    Mismatches: 148
Query Match: 5.25%             Indels: 158
DB: 1                       Gaps: 24

US-09-147-052-1 (1-1371) x PCT-US02-37235-44 (1-1531)

QY 43 ATCTATATGTTAGTACGACTCATCTCCGAGTACCCAAATGTG----- 84
Db 329 IleSerLeuGlyAlaAsnAlaLysAlaSerGlnAspAsnValValAlaIleGlyLysTyr 348
85 ---ACATCAAGAGAGTGTGTTGCGAGCGTCCAGTGTGCTGAGGAAGAGTCTACGTTT--- 138
Db 349 AlaThrAlaThrGluSerGlySerMetAlaIleGlyGlnGlyAlaLysSerThrPheLys 368
139 TATCTTTGTCCTCCACCAGCGGTTCACCGTGATCGGTGATAGTTCGCGCTGTATGTCT 198
Db 369 AsnSerLeuAlaLeuGlyThrGlyThrIleValAsnSerValAspGlyGlnSerLys 388
199 ATTACTAAAAAGATGCAACCCCAATAATATGCGCAACCCCAATAATAGACGCGCGAATG 258
Db 389 PheThrAlaGlnAsnTyrAspAlaAsnAsnGlyValValAlaAlaAsnAlaGlyLys 408
259 GAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCC 318
Db 409 Glu---ArgArgIleIleAsn----- 414
319 AAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAACACGTTTACAAATACCTT 378
Db 415 -----ValAlaGlyGlyArgAsnAspThrAspAlaValAsn----- 426
379 AATGCAACATTAGAACAACTAAATATGCTAAACCTAATTTAGAAATCAGCCATC----- 432
Db 427 -----IleAlaGlnLeuLysPheValAsnAspAsnLeuAlaLysSerIleAlaGly 443
433 -----AACCAAGCTAATACGGATAAAGCACT-----TTTGATAAT 468
Db 444 AlaGlyTyrAsnGlyTyrGluThrAspGlyHisThrTyrLysAlaProValPheSerIle 463
469 GAACACCCCAATTTAGTTGAAGCATACAAAGCACTAAACCACTTAGAACACGCTGT 528
Db 464 LysAsnThrAsn-----TyrHisAspValLysThrAlaValGluAlaAlaGln 479
529 ACTAACCTTGAAGTTGTTCATCACTGCTTATATCAATTCGAATTCGCAATTAATTAGTGGAT 588
Db 480 ThrAsnTyrValSerValAsnSerThr----- 488
589 CTATACAAATAGCTAGTAGTTTAACTAACTAAACACTAGATCCACTAAATGGGGGAACG 648
Db 489 -----AsnThr 490
649 CTTTATAGATTCTAATGAGATTACTACAGCTAATAAGATATTAATAATACGTTATCACT 708
Db 491 AlaAlaAspSerAsnTyrAspAsnLysGlyAlaLysAlaValGlySerIle-----Ala 508
709 ATTATGACAAAGACCTAATGCTGATCATATCAATAGTTTATTAAAAAGTGATT 768
Db 509 LeuGlyGluLysAlaThrThrGlyThrAlaAlaMetAsnSer----- 522
769 CAAATAATGAACAAAGTTTGTGGGACTTTTACAAACGCTAATGTTCAACCTTCAAC 828

```

```

Db 523 -----IleAlaIleGlyLeuAsnSerAsnVal-----SerGly 533
QY 829 TACAGTTTGTGCTTTTGTAGTGTGATGTAACACCCGCTCAATTATAATATGCAAGAAG 888
Db 534 GlnAsnThrValAlaLeuGlyAlaAsnIleThrAlaThrThr----- 547
QY 889 ACCGTTTGGGAATGCTGATGAACCTTCAAGTAGAATCTTCCAAACACGCAATAGTATCACA 948
Db 548 -----AsnGly-----SerValIleLeuGlyAsnSerSerThr----- 558
QY 949 GATGTTTCTTGGGAATTATAGTTTGTAGTGGAAACACACGAGGATACCAATTTAGTTTTC 1008
Db 559 -----ThrGluGlySerHisProValSer 566
QY 1009 AACTATGTCCTCACT-----GGTATTATAT-----TTCCCTTATAAGTTGGTT 1056
Db 567 AsnValSerSerAlaThrValAsnGlyTyrThrTyrSerGlyPheThrGlyThrValLys 586
QY 1057 AAAGCAGCTGATGCTAATAACGTTGGATTACAATAACAATAATAATGGAATGTTCAA 1116
Db 587 GluSerGlyHisPheValSerIleGly-----SerLysGlyAsnGluArg 601
QY 1117 CAAGTTGAGTTTCCCACTTCAACTAGTGTGCAAAATTAATACAGTCAATCCACCTCCAGCA 1176
Db 602 GlnIleLys-----AsnValAlaAlaGlyAsnValAlaAlaAsnSerThrAspAla 618
QY 1177 GTTGATGAGATTAAAGCTGCTAAATC-----GTTTAA 1209
Db 619 ValAsnGlySerGlnLeuPheAlaValAlaSerArgValGluGlnGlyTrpGlnIleThr 638
QY 1210 TCAGTTTAAAGATTGTC-----CAAAAC-----ACAATCGAATTA 1245
Db 639 SerGlyValGluAsnGlyGlyThrGlnAsnGlyAlaAlaSerThrAlaThrIleLysPro 658
QY 1245 ACTGTTCCACCGGTCGACGAATAATGATTAAGTTGCCCAATGATGTC-----AACATT 1302
Db 659 Ser-----AsnGlnValLysLeuLeuAlaGlyLysAsnLeu 670
QY 1303 TATCTTAGCTCAATGAAATAAT 1326
Db 671 AlaValLysGlnAsnGlyThrAsn 678

RESULT 14
PCT-US03-20460-14
; Sequence 14, Application PC/TUS0320460
; GENERAL INFORMATION:
; APPLICANT: Iowa State University Research Foundation, Inc. and Department of
; APPLICANT: Agriculture for
; APPLICANT: and on behalf of the State of New South Wales
; TITLE OF INVENTION: Immunogenic Mycoplasma Hypopneumoniae
; TITLE OF INVENTION: Polyptides
; FILE REFERENCE: 08411/035W01
; CURRENT FILING DATE: 2003-06-27
; PRIOR FILING DATE: 2003-06-27
; PRIOR FILING DATE: 2003-06-27
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 945
; TYPE: PRT
; ORGANISM: Mycoplasma hypopneumoniae
PCT-US03-20460-14

Alignment Scores:
Pred. No.: 0.104      Length: 945
Score: 121.00         Matches: 98
Percent Similarity: 33.60%      Conservative: 69
Best Local Similarity: 19.72%    Mismatches: 164
Query Match: 5.20%             Indels: 166
DB: 1                       Gaps: 21

```

QY	1006	AGCAACTATGGTCCACCTAAGTTATTTATATTTCCTTTAAAGTTGGTTAAAGCAGCT	1065
	:::	:::	
Db	635	GluAsp-----AspLysLeuValPheAspPheLysLeuValSerLysAla	649
	:	:::	
QY	1066	GAT-----GCTAATAACGTTGGATTACAATACAAATAAATAATGGA	1107
		::: ::: :	
Db	650	AspArgAlaIleIleThrGlnGlySerLysIleSerLeuGlyLeuIleAsnSerAsp	669
		::: ::: :	
QY	1107	-----	1107
Db	670	LysSerAlaThrAspGluIleLysGlnPheSerProAspLeuPheLeuAspAlaThrIle	689
	:	:::	
QY	1108	-----AATGTTCAACAAGTTTGAGTTGCCACTTCAACTACTGCA	1146
		: ::::: :	
Db	690	GlyGluGlnSerAspTyrLysAsnLysGlnLysAspTyrThrLeuLysSerLeuArg	709
	:	:::	
QY	1147	AAT-----AATACTACAGCTTAATCCAACCTCCAGCA	1176
	:::	:::	
Db	710	AspLeuMetGlyAsnGlyPheValTyrLysProGluThrLysSerAsnProGlnGluAsn	729
	:	:::	
QY	1177	GTTGATGAGATTAAAGTTGCT-----AAAATCGTTTTATCAGGTTTAAGATTGGC	1227
		: :	
Db	730	ValLeuLysLeuGlnThrGlySerGluGlnLysProLeuProGlyLeuArgSerGly	749
	:	:::	
QY	1228	CAAACACATCATTAAGTGTTCACCGGTGAAGGAATATGAATAA	1278
		: :	
Db	750	LeuIleTyrIleAlaPheThrVal-----AsnAsnIleAsnLys	762
	:	:::	
RESULT 15			
US-09-200-650E-5			
; Sequence 5, Application US/09200650E			
; GENERAL INFORMATION:			
; APPLICANT: Patti, Joseph M.			
; APPLICANT: Foster, Timothy J.			
; APPLICANT: Hook, Magnus A.O.			
; APPLICANT: Eidhinn, Deirdre N1			
; APPLICANT: Perkins, Samuel L.			
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus			
; FILE REFERENCE: P06283US2/BAS			
; CURRENT APPLICATION NUMBER: US/09/200,650E			
; CURRENT FILING DATE: 1998-11-25			
; PRIOR APPLICATION NUMBER: 60/066,815			
; PRIOR FILING DATE: 1997-11-26			
; PRIOR APPLICATION NUMBER: 60/098,427			
; PRIOR FILING DATE: 1998-08-31			
; NUMBER OF SEQ ID NOS: 23			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 5			
; LENGTH: 1315			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-09-200-650E-5			
Alignment Scores:			
Pred. No.:	0.124	Length:	1315
Score:	120.00	Matches:	119
Percent Similarity:	36.82%	Conservative:	78
Best Local Similarity:	22.24%	Mismatches:	202
Query Match:	5.16%	Indels:	136
DB:	5	Gaps:	28
US-09-147-052-1 (1-1371) x US-09-200-650E-5 (1-1315)			
QY	58	AACATCTCCGAGTACCACCAATGTGCATCAAGAGAGTCTTTCCGAGCGTC	111
		:::	
Db	82	AsnGlnGluAspAsnThrLysAsnAspAsnGlnLysGluMetValSerSerGlnGlyAsn	101
	:	:::	
QY	112	-----CAGTGTCTGGAAGAGCTCTACGTTTATCTTTGTCCC	150
		:	
Db	102	GluThrThrSerAsnGlyAsnLysLysLysLysGluSerValGlnSerThrThrGly	121
	:	:::	
QY	151	CCACCGAGTGGT---TCAACCGGTGATCCCGCTAGAATTCGGCTGTATGCTATTACTAAA	207

```

Db      122 AsnLysValGluValSerThrAlaLysSerAspGluGlnAlaSerProLysSerThrAsn 141
QY      208 AAAGATGCAACCCCAATAATAGCGCAACCCCAATAGAGCAGCGCGAATGGAGTTAACA 267
Db      142 GluAspLeuAsnThrLysGlnThrIleSerAsnGlnGluAlaLeuGlnProAspLeuGln 161
QY      268 GAT-----CTAATCAATGCTAAACCGATGACATAGCTTCCACACAGACTATGCC 318
Db      162 GluAsnLysSerValValAsnValGlnProThrAsnGluGluAsnLysLysValAspAla 181
QY      319 AAGATTGAAGCTAGCTTATCATCTGCTTATAGTGAAGCT-----GAAACA 363
Db      182 LysThrGluSerThrThrLeuAsnValLysSerAspAlaLysLysSerAsnAspGluThr 201
QY      364 ---GTTAAACAATACCTTAATGCAACATTAGAA-----CAACTAAATAATGGCTAAA 411
Db      202 LeuValAspAsnAsnSerAsnSerAsnGluAsnAlaAspIleIleLeuProLys 221
QY      412 ACTAATTTAGAAATCAGCCATCAAC-----435
Db      222 SerThrAlaProLysArgLeuAsnThrArgMetargIleAlaValGlnProSerSer 241
QY      436 -----CAAGCTAATACGGATAAAACGACTTTTGAT 465
Db      242 ThrGluAlaLysAsnValAsnAspLeuIleThrSerAsnThrThrLeuThrValValAsp 261
QY      466 AATGAACACCCCAATTTAGTT-----GAAGCATACAAAGCAGCTAAACCACTTTA 516
Db      262 AlaAspLysAsnAsnLysIleValProAlaGlnAspTyrLeuSerLeuLysSerGlnIle 281
QY      517 GAACAACGCTCTACTAACCTTGAAGGTTTGTCTCACTCACTGTTATATCAATCAATTCG--- 573
Db      282 -----ThrValAspAspLysValLysSerGlyAspTyrPheThrIleLysTyr 297
QY      574 AATAATTTAGTGTACTATACAAAT-----AAGCTAGTAGTGTATTAATACTAAACA 624
Db      298 SerAspThrValGlnValTyrGlyLeuAsnProGluAspIleLysAsnIleGlyAspIle 317
QY      625 CTAGATCCCAATAATGGGGAAGCTTTAGATTCTAATGAGATTACTACAGCTAATAAG 684
Db      318 LysAspProAsnAsnGlyGluThrIleAlaThrAlaLysHis---AspThrAlaAsnAsn 336
QY      685 AATATTAAATAACGTTATCACTATTAATGAACAAACAGCACTAATGCTGATGCTATCT 744
Db      337 LeuIleThrTyrThrPheThrAspTyrValAspArgPheAsnSerValGlnMetGlyIle 356
QY      745 AATAGTTTATTATAAAAA-----GTGATTCAAATAATGAACAAAGTTT 789
Db      357 AsnTyrSerIleTyrMetAspAlaAspThrIleProValSerLysAsnAspValGluPhe 376
QY      790 -----GTAGGG-----ACTTTTACAACGCTAATGTTCAACCTTCAAACTAC 831
Db      377 AsnValThrIleGlyAsnThrThrLysThrThrAlaAsnIleGlnTyrProAspTyr 396
QY      832 AGTTT-----GTTGCTTTTAGTGTGATGATAACACCCGCTCAAT 870
Db      397 ValValAsnGluLysAsnSerIleGlySerAlaPheThrGluThrValSerHisValGly 416
QY      871 TATAAA-----TATGCAAGAGCGCTTTGGAATGGTGATCAACCTTCAAGT 918
Db      417 AsnLysGluAsnProGlyTyrTyrLysGlnThrIleTyr-----429
QY      919 AGAATTCCTTGCACACGAAATAGTATCACAGATGTTCTTGGATTATAGTTAGTGGGA 978
Db      430 ---ValAsnProSerGluAsnSerLeuThrAsn-----439
QY      979 ACAAAACCAAGTACCAATTTAGTTTACCAACTGTCATCACTGGTTATTATA--- 1035
Db      440 AlaLysLeuLysValGlnAlaTyrHisSerSerTyrProAsnAsnIleGlyGlnIleAsn 459
QY      1036 -----TATTTCCCT-----TATAAGTTGGTTAAAGCA 1062

```

Search completed: October 8, 2003, 19:16:25
 Job time : 45.327 secs

```

Db      460 LysAspValThrAspIleLysIleTyrGlnValProLysGlyTyrThrLeuAsnLysGly 479
QY      1063 GCTGATGCTAAT-----AACGTTGGATTACAAATAC-----AAATTTAAAT 1101
Db      480 TyrAspValAlaAsnThrLysGluLeuThrAspValThrAsnGlnTyrLeuGlnLysIleThr 499
QY      1102 AATGGA-----AATGTTCAACAAAGTTGAGTTTCCCACTTCAACTAGTGCATAAT----- 1149
Db      500 TyrGlyAspAsnAsnSerAlaValIleAspPheGlyAsnAlaAspSerAlaTyrValVal 519
QY      1150 -----AATACTACAGCTAATCCCACTCCAGCAGCTTGCATGAGATTAAAGTTGTAAATC 1203
Db      520 MetValAlaAsnThrLysPheGlnTyrThrAsnSerGluSerProThrLeuValGlnMetAla 539
QY      1204 GTTTTATCAGGTTTAAAGATTGGCCAAACACACATCGAATTAAGTTGTCCCAACGGGT--- 1260
Db      540 ThrLeuSer-----SerThrGlyAsnLysSerValSerThrGlyAsn 553
QY      1261 -----GAAGGAAATATGAATAAAGTTGCGCCAATGATTGGC 1296
Db      554 AlaLeuGlyPheThrAsnAsnGlnSerGlyGlyAlaGlyGlnGluValTyrLysIleGly 573
QY      1297 AAC---ATTTATCTTAGCTCAATGAATAATAAGTGTGATAAGATC 1338
Db      574 AsnTyrValTrpGluAspThrAsnLysAsnGlyValGlnGluLeu 588

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:53:53 ; Search time 414.851 Seconds
(without alignments)
6014.205 Million cell updates/sec

Title: US-09-147-052-1
Perfect score: 2326
Sequence: 1 atgcaatttttagcgga.....gacccggtagacattttataa 1371

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5728757 seqs, 90918778 residues

Total number of hits satisfying chosen parameters: 11457514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Q-/cgn2_1/USPTO_spool/US09147052/runat_08102003_154343_29901/app_query.fasta_1.5980
-DB-Pending_Patents_AA_Main -QFWT-fastan -SUFFIX-rapm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09147052@cgn_1.1.2602/runat_08102003_154343_29901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_AA_Main:
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2324	99.9	456	15	US-09-147-052-2
2	2317	99.6	456	24	US-09-901-572A-3
3	2015	86.6	1086	15	US-09-147-052-4
4	1929.5	83.0	384	27	US-10-131-591A-12
5	1777	76.4	357	27	US-10-131-591A-23
6	750.5	32.3	645	22	US-09-791-537-105581
7	749	32.2	650	22	US-09-791-537-116408
8	715.5	30.8	647	6	US-08-014-631-7
9	715.5	30.8	647	6	US-08-230-312-7
10	715.5	30.8	647	22	US-09-791-537-7385
11	715.5	30.8	647	22	US-09-791-537-34555
12	320	13.8	62	27	US-10-131-591A-5
13	314	13.5	865	3	US-07-722-860-13
14	314	13.5	865	6	US-08-213-449A-13
15	314	13.5	865	6	US-08-213-449B-13
16	308	13.2	62	27	US-10-131-591A-6
17	187.5	8.1	1095	21	US-09-710-279-3154
18	185.5	8.0	6641	28	US-10-282-122A-70580
19	185.5	8.0	10182	26	US-10-092-411A-3159
20	185.5	8.0	10203	18	US-09-450-969-4098
21	177	7.6	1347	12	US-08-827-356-5566
22	177	7.6	1347	20	US-09-611-529-4400
23	177	7.6	1347	25	US-09-950-084-4400
24	176.5	7.6	6713	28	US-10-282-122A-43811
25	176.5	7.6	1291	32	US-60-257-931-3393
26	176.5	7.6	1291	32	US-60-269-308-4414
27	176.5	7.6	2398	28	US-10-282-122A-70176
28	176.5	7.6	6281	1	PCT-US02-03987-12996
29	176.5	7.6	6281	23	US-09-815-242-12996
30	176.5	7.6	6281	26	US-10-072-851-12996
31	175	7.5	3351	12	US-08-827-356-5564
32	175	7.5	3351	20	US-09-611-529-4562
33	175	7.5	3351	25	US-09-950-084-4562
34	174	7.5	2086	1	PCT-US02-03987-5639
35	174	7.5	2086	23	US-09-815-242-5639
36	174	7.5	2086	26	US-10-072-851-5639
37	174	7.5	2086	32	US-60-242-578-948
38	174	7.5	2086	32	US-60-253-625-2292
39	174	7.5	2086	32	US-60-257-931-3398
40	174	7.5	2086	32	US-60-269-308-4231
41	174	7.5	5795	1	PCT-US02-03987-12610
42	174	7.5	5795	23	US-09-815-242-12610
43	173.5	7.5	496	26	US-10-072-851-12610
44	173.5	7.5	496	29	US-10-311-879-29
45	173.5	7.5	933	14	US-09-095-272-2

ALIGNMENTS

RESULT 1

US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167

; CURRENT APPLICATION NUMBER: US/09/147,052
 ; CURRENT FILING DATE: 1999-04-05
 ; PRIOR APPLICATION NUMBER: JP 08-103548
 ; PRIOR FILING DATE: 1996-03-29
 ; PRIOR APPLICATION NUMBER: PCT/JP97/01084
 ; PRIOR FILING DATE: 1997-03-28
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: hybrid
 US-09-147-052-2

Alignment Scores:

Pred. No.: 7,69e-201 Length: 456
 Score: 2324.00 Matches: 456
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 15 Gaps: 0

US-09-147-052-1 (1-1371) x US-09-147-052-2 (1-456)

QY 1 ATGCACATATTTAGCGGCAATTCATATTTTCCCTTATAGTATTCCTATGCTACGCAAC 60
 DB 1 MethIstYrPheArGArGsnCysIlePhePheLeuIleValIleLeuYrGlyThrAsn 20
 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTGCGAGCGCCAGTTGCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGCTAGCTTTATCTTTGTCCTCCACAGTGGGTCAACCGGTGATCCGCTA 180
 DB 41 GluGluGluSerThrPheYrLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAATTCGGCTGTATGCTATTACTTAAAGATGCAACCAATATGCAACCAACCA 240
 DB 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
 QY 241 TTAGAAGCAGCGGAGTGGAGTAAACAGATCTAATCAATGCTAAAGCGATGACATTAGCT 300
 DB 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
 QY 301 TCACTCAAGACTATGCAAGATGAGCTAGTATTATCATCTGCTTATAGTGAAGCTGAA 360
 DB 101 SerLeuGlnAspYrAlaLysIleGluAlaSerLeuSerSerAlaYrSerGluAlaGlu 120
 QY 361 ACAGTTAACAAATACCTTAATGCAACATTAAGAACAACTAAATAATGGCTAAACCTAATTTA 420
 DB 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
 QY 421 GAATCAGCCATCAACAGCTAATACGATAAAGACTTTTGTATGATGAACACCCAAAT 480
 DB 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
 QY 481 TTAGTTGAAGCATACAAAGCACTAAACACCTTTAGAACAGCTCTACTACTACCTTGA 540
 DB 161 LeuValGluAlaYrLysAlaLeuLysThrLeuGluGlnArgAlaThrAsnLeuGlu 180
 QY 541 GGTTCGTCACTCACTTATATCAAAATTCGCAATAATTTAGTGGATCTATACAATAA 600
 DB 181 GlyLeuSerSerThrAlaYrAsnGlnIleArgAsnAsnLeuValAspLeuYrAsnLys 200
 QY 601 GCTAGTATTATTAATAACACATAGATCCACTAAATGGGGAACGCTTTTAGATTCT 660
 DB 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
 QY 661 AATGAGATTACTACAGTAAAGATATTAATACCTTATCACTATTATTAATGACAA 720
 DB 221 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
 QY 721 AAGACTAAATGCTGATGATCTATCTAATAGTATTATTAATAAAGATGATCAAAATAATGAA 780

DB 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGlu 260
 QY 781 CAAGTATTGCTAGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTGT 840
 DB 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnYrSerPheVal 280
 QY 841 GCTTTTAGTCTGATGAACACCGCTCAATATATAATATCAAGACGACCGTTTGAAT 900
 DB 281 AlaPheSerAlaAspValThrProValAsnYrLysYrAlaArgThrValTrpAsn 300
 QY 901 GGTGATGAACCTTCAAGTAGAATTTCTGCAACACGATATGATCATCAGATGTTCTTGG 960
 DB 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 320
 QY 961 ATTTATAGTTAGCTGGAACAAACACAGTAGTACCAATTTAGTTTACCACTATGTCCTCA 1020
 DB 321 IleYrSerLeuAlaGlyThrAsnThrLysYrGlnPheSerPheSerAsnYrGlyPro 340
 QY 1021 TCAACTGGTTATTTATTTCCCTTATAAGTTGTTAAAGCAGCTGATGCTAATAACGTT 1080
 DB 341 SerThrGlyYrLeuYrPheProYrLysLeuValLysAlaAlaAspAlaAsnVal 360
 QY 1081 GGATTCAATACAAATTAATATGAAATGTTTCAACAAGTTGAGTTGGCACTTCAACT 1140
 DB 361 GlyLeuGlnYrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
 QY 1141 AGTCAAAATATATCTACAGCTATCCCAACTCCAGCAGCTTGTGATGAGATTAAGTTGCTAAA 1200
 DB 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
 QY 1201 ATCGTTTATCAGTTTAAGATTGGCCAAACACAACTCGAATTAAGTTTCCAAACGGT 1260
 DB 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
 QY 1261 GAAGGAATATGAATAAGTTGCGCCCAATGATGGCAACATCGAATTAAGTTTCCAAACGGT 1320
 DB 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleYrLeuSerSerAsnGlu 440
 QY 1321 AATAATGCTGATAGATCCCGGTACCGTCCGACCGGTGATGATTTTAA 1368
 DB 441 AsnAsnAlaAspLysIleProGlyYrArgArgProGlyThrPheLeu 456

RESULT 2

US-09-901-572A-3
 ; Sequence 3, Application US/09901572A
 ; GENERAL INFORMATION:
 ; APPLICANT: Nippon Zeon Co., Ltd.,
 ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
 ; FILE REFERENCE: J209
 ; CURRENT APPLICATION NUMBER: US/09/901,572A
 ; CURRENT FILING DATE: 2003-03-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: Mycoplasma gallisepticum
 ; FEATURE:
 ; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
 US-09-901-572A-3

Alignment Scores:

Pred. No.: 3.3e-200 Length: 456
 Score: 2317.00 Matches: 454
 Percent Similarity: 99.78% Conservativeness: 1
 Best Local Similarity: 99.56% Mismatches: 1
 Query Match: 99.61% Indels: 0
 DB: 24 Gaps: 0

US-09-147-052-1 (1-1371) x US-09-901-572A-3 (1-456)

QY 1 ATGCACATATTTAGCGGCAATTCATATTTTCCCTTATAGTATTCCTATGATGATGACGAAC 60

Db	1	MetHisTyrPheArgArgAsnGlnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn	20
Qy	61	TCATCTCCGAGTACCCAAAATGTGACATCAAGAGAGAGTGTCTTTCGAGCGCTCCAGTTGTCT	120
Db	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40
Qy	121	GAGAAAGAGCTACGTTTATCTTTGTCTCCCCACCAGTGGGTTCACACCGTGATCCGCTCA	180
Db	41	GluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu	60
Qy	181	GAATCGCGTGTATGCTATTAACAAAAGATGCAACCCCAATAATATGCCAAACCCAA	240
Db	61	GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln	80
Qy	241	TTAGAACACGCGGAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCT	300
Db	81	LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla	100
Qy	301	TCACTACAGACATATGCCAAGATGAAGCTAGTGTATCATCTGCTATAGTGAAGCTGAA	360
Db	101	SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu	120
Qy	361	ACAGTTAACCAATACCTTAATGCAACATTTAGACACACTAAAAATGGCTAAACTAATTTA	420
Db	121	ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu	140
Qy	421	GAATCAGCCATCAACCAAGCTAATACGGATAAAGCACTTTTCATATGTAATGAACACCCAAAT	480
Db	141	GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn	160
Qy	481	TTAGTTGAAGCATACAAAGCACTAAACACACTTTTAGAACACGCTGCTACTACCTTGAA	540
Db	161	LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu	180
Qy	541	GGTTTGTCATCACTGCTTATATCAAAATTCGCAATAATTTAGTGGATCTATACATAAAA	600
Db	181	GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys	200
Qy	601	GCTAGTAGTTTAATACTAAACACTAGATCCACTAATGGGGAACGCTTTTAGATTCT	660
Db	201	AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer	220
Qy	661	AATGAGATTACTACAGCTAATAGAATTAATAATACGTTATCAACTATTAATGAACAA	720
Db	221	AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln	240
Qy	721	AAGACTAATGCTGATGATTCATCTAATAGTTTATTAAAAAGTGATTCAAATATATGAA	780
Db	241	LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGlu	260
Qy	781	CAAAAGTTTGTAGGAGCTTTTACAAACGCTTAATGTTCAACGCTTCAAACTACAGTTTGT	840
Db	261	GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal	280
Qy	841	GCTTTTAGTGCTGATGTAAACACCCGTCAAATATAAATATGCAAGAGGACCGTTTGGAA	900
Db	281	AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn	300
Qy	901	GGTGATGAACCTTCAAGTAGAATCTTGCAACACAGAAATAGTATCACAGATGTTCTGTG	960
Db	301	GlyAspGluProSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp	320
Qy	961	ATTATATAGTTTACGTGGAACAAACACCAAGTACCAATTTAGTTTACCACTATGGTCCA	1020
Db	321	IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro	340
Qy	1021	TCAACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAACAGCTGATGCTAATAAGTT	1080
Db	341	SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnAsnVal	360
Qy	1081	GGATTACAATAAATAATATGGAATGTTTCAACAAGTTGAGTTTGGCTTCAACT	1140

Db	361	GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr	380
Qy	1141	AGTGCAAAATAACTACAGCTAATCCAACTCCAGACGTTGATGAGATTAAAGTTGCTAAA	1200
Db	381	SerAlaAsnAsnThrThrAlaAsnProThrProAlaValaspGluIleLysValAlaLys	400
Qy	1201	ATCGTTTTATCAGGTTTAAAGATTGGCCAAACAAATCGAAATTAAGTTGTCCAACGGGT	1260
Db	401	IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly	420
Qy	1261	GAGGAAATATGAATAAGTTGGCCCAATGATTCGCAACATTTATCTTACCTCAAAATGAA	1320
Db	421	GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu	440
Qy	1321	ATAAATGCTGATAGATCCCGGGTACCGTCGACCGGTACATTTTAA 1368	
Db	441	AsnAsnAlaaspLysIleProGlyTyrArgArgProGlyThrPheLeu 456	
RESULT 3			
US-09-147-052-4			
; Sequence 4, Application US/09147052			
; GENERAL INFORMATION:			
; APPLICANT: SAITOH, Shuji			
; APPLICANT: TSUZAKI, Yoshinari			
; APPLICANT: YANAGIDA, Noboru			
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,			
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE			
; FILE REFERENCE: 981167			
; CURRENT APPLICATION NUMBER: US/09/147,052			
; CURRENT FILING DATE: 1999-04-05			
; PRIOR APPLICATION NUMBER: JP 08-103548			
; PRIOR FILING DATE: 1996-03-29			
; PRIOR APPLICATION NUMBER: PCT/JP97/01084			
; PRIOR FILING DATE: 1997-03-28			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 4			
; LENGTH: 1086			
; TYPE: PRT			
; ORGANISM: hybrid			
US-09-147-052-4			
Alignment Scores:			
Pred. NO.:	7.9e-173	Length:	1086
Score:	2015.00	Matches:	396
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	86.63%	Indels:	0
DB:	15	Gaps:	0
US-09-147-052-1 (1-1371) x US-09-147-052-4 (1-1086)			
Qy	181	GAATTCGGCTGTATCTCTATTACTAAAAAGATGCAACCCCAATAATGGCCAAACCCAA	240
Db	691	GluPheGlyCysMetSerIleThrLysLysaspAlaAsnProAsnAsnGlyGlnThrGln	710
Qy	241	TTAGAAGCAGCGCAAGTGGAGCTTAACAGATCTCAATCAATGCTAAGCGATGACATTAGCT	300
Db	711	LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla	730
Qy	301	TCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA	360
Db	731	SerLeuGlnaspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu	750
Qy	361	ACAGTTAACATAACCTTAATGCAACATTAGAACAACATAAAATGGCTAAAACTAAATTTA	420
Db	751	ThrValAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu	770
Qy	421	GAATCACCCTCAACCAAGCTAATACCGATAAAGACACTTTTGATATGAACACCCCAAT	480
Db	771	GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn	790
Qy	481	TTAGTTGAAGCATCAAAAGCACTAAAAACCACTTTTAGAACAAACGCTGCTACTAACTTGA	540

```
Db 791 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 810
Qy 541 GCTTTGTCATCACTGCTATAATCAAAATCGCAATATTTAGTGGATCTATACAAATAA 600
Db 811 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 830
Qy 601 GCTAGTATTAACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCT 660
Db 831 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 850
Qy 661 AATGAGATTACTACAGCTAATAAGAAATTAATATACGCTTATCAACTATTAAATGAACA 720
Db 851 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 870
Qy 721 AAGACTAATGCTGATGCTATCTAATTAATAGTTTATAAAGATGATTCAAAATAATGAA 780
Db 871 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnGlu 890
Qy 781 CAAAGTTTGTAGGACCTTTTACAAACGCTTAATGTTCAACCTTCAAACTACAGTTTGT 840
Db 891 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 910
Qy 841 GCTTTTAGTGTGATGTAACACCGCTCAATTAATAATATGCAAGACCGTTTGGAT 900
Db 911 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn 930
Qy 901 GGTGATGAACCTTCAAGTAGAATCTTGCACAAACGAGTACTATCACAGATGTTCTTGG 960
Db 931 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 950
Qy 961 ATTTATAGTTAGTGGAAACAAACACGAGTACCAATTTAGTTTAGCAACTATGGTCCA 1020
Db 951 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 970
Qy 1021 TCACTGGTTATTATATTTCCCTATTAAGTTGTTAAAGCAGCTGATGCTTAATAACGTT 1080
Db 971 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 990
Qy 1081 GGATTACAATAACAATAATGGAATGTTCAACAAGTTGAGTTTGCACCTTCAACT 1140
Db 991 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 1010
Qy 1141 AGTGAATAATATCTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAA 1200
Db 1011 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 1030
Qy 1201 ATCGTTTTATCAGGTTTAAGATTGGCCAAACACAATCGAATTAAGTTCCCAACGGGT 1260
Db 1031 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 1050
Qy 1261 GAAGGAATATCAATAAAGTTGCGCAATGATTGGCAACATTTATCTTTAGCTCAAAATGAA 1320
Db 1051 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 1070
Qy 1321 AATATGCTGATAGATCCCGGGTACCGTCCGACCGCGGTACATTTTA 1368
Db 1071 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 1086
```

RESULT 4

US-10-131-591A-12

; Sequence 12, Application US/10131591A

; GENERAL INFORMATION:

; APPLICANT: Nippon Zeon Co., Ltd.,

; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof

; FILE REFERENCE: J209

; CURRENT APPLICATION NUMBER: US/10/131,591A

; CURRENT FILING DATE: 2002-08-15

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12

; LENGTH: 384

; TYPE: PR

```
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12
```

```
Alignment Scores:
Pred. No.: 3,41e-165 Length: 384
Score: 1929.50 Matches: 382
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.22% Mismatches: 1
Query Match: 82.95% Indels: 1
Db: 27 Gaps: 1
```

US-09-147-052-1 (1-1371) x US-10-131-591A-12 (1-384)

```
Qy 181 GAATTCGGCTCTATGCTCTATTACTAAAAAGATCAAAACCCCAATAATGGCCAAACCCAA 240
Db 1 GluPhe---CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 19
Qy 241 TTAGAAGCAGCGCAATGGAGTTAACAGATCAATCAATGCTAAAGCGCATGACATTAGCT 300
Db 20 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 39
Qy 301 TCACCTACAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTAGAAGCTGAA 360
Db 40 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 59
Qy 361 ACAGTTAAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAACATAATTTA 420
Db 60 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 79
Qy 421 GAATCAGCCATCAACCAAGCTAATACGAGTAAACGACTTTTGTATTAATGAACCCCAAT 480
Db 80 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 99
Qy 481 TTAGTTGAGCATACAAGCAGCTAAAAACCCACTTTAGAACACGCTGCTACTAACTTGA 540
Db 100 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 119
Qy 541 GGTTTGTTCATCACTGCTTATAATCAAAATTCGCAATAATTTAGTGGATCTATACAAATAA 600
Db 120 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 139
Qy 601 GCTAGTATTAACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCT 660
Db 140 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 159
Qy 661 AATGAGATTACTACAGCTAATAAGAAATTAATATACGCTTATCAACTATTAAATGAACA 720
Db 160 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 179
Qy 721 AAGACTAATGCTGATGCTATCTAATAGTTTATAAAGATGATTCAAAATAATGAA 780
Db 180 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnGlu 199
Qy 781 CAAAGTTTGTAGGACCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGT 840
Db 200 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 219
Qy 841 GCTTTTAGTGTGATGTAACACCGCTCAATTAATAATATGCAAGACCGTTTGGAT 900
Db 220 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTrpAsn 239
Qy 901 GGTGATGAACCTTCAAGTAGAATCTTGCACAAACGAGTATGATCACAGATGTTCTTGG 960
Db 240 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 259
Qy 961 ATTTATAGTTAGTGGAAACAAACACGAGTACCAATTTAGTTTAGCAACTATGGTCCA 1020
Db 260 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 279
Qy 1021 TCACTGGTTATTATATTTCCCTATTAAGTTGTTAAAGCAGCTGATGCTTAATAACGTT 1080
```

Db 280 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 299
QY 1081 GGATTACATCAATTAATAATGGAATGTCACAAAGTTGAGTTGGCCACTTCAACT 1140
Db 300 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 319
QY 1141 AGTCAAAATATACACTACAGCTTAATCCAACTCCACAGCTGATGAGATTAAAGTTGCTAAA 1200
Db 320 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluLeuLysValAlaLys 339
QY 1201 ATCGTTTATCAGTTTAAAGATTGGCCAAACAAATCGAATTAAGTTTCCACGGGT 1260
Db 340 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 359
QY 1261 GAAGCAATATGAATAAGTTGCCCAATGATGGCAACATTTATCTAGTCAATGAA 1320
Db 360 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 379
QY 1321 AATAATGCTGATAAG 1335
Db 380 AsnAsnAlaAspLys 384

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified T71-1 portion (downstream of BglI) of
; OTHER INFORMATION: PN240K-S
US-10-131-591A-23

Alignment Scores:
Pred. No.: 2 05e-151 Length: 357
Score: 1777.00 Matches: 352
Percent Similarity: 98.60% Conservative: 0
Best Local Similarity: 98.60% Mismatches: 5
Query Match: 76.40% Indels: 0
DB: 27 Gaps: 0

US-09-147-052-1 (1-1371) x US-10-131-591A-23 (1-357)

QY 265 ACAGATCTAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAGACTATGCCAAGATT 324
Db 1 ThrAspLeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIle 20
QY 325 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTACATTAACCTTAATGCA 384
Db 21 GluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValaAsnAsnLeuGlnAla 40
QY 385 ACATTAGAACAACATAAAATGGCTAAACTAATTTAGAAATCAGCCATCAACCAAGCTAAT 444
Db 41 ThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsn 60
QY 445 ACGGATAAAGCAGCTTTTGTATAATGAACCCCAATTTAGTTGAAGCATACAAAGCAGCTA 504
Db 61 ThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeu 80
QY 505 AAAACCACTTTAGAACACAGCTGCTACTTAACCTTGAAGTTTGTTCATCACTGCTTATAAT 564
Db 81 LysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsn 100
QY 565 CAATTCGCAATATTTAGTGGATCTATACATAAAGCTAGTAGTTTAATAACTAAAAACA 624

Db 101 GlnIleArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThr 120
QY 625 CTAGATCCACTAAATGGGGAAACGCTTTTAGATTCTTAATGAGATTACTACAGTAATAAG 684
Db 121 LeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLys 140
QY 685 AATATTAAATACAGTTTATCAACTATTAAATGACAAAAGACTAATGCTGATGCAATTACT 744
Db 141 AsnIleGlnAsnThrLeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSer 160
QY 745 AATAGTTTATTAATAAAGTCAATCAAAATAATGAACAAAGCTTTTGTAGGACATTTTACA 804
Db 161 AsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThr 180
QY 805 AACGCTAATGTTCACCTTCAAACTTCAAACTACAGATTTTGTGTGTTTAGTGTGATGTAACACCC 864
Db 181 AsnAlaAsnValGlnProSerGlnTyrSerPheValAlaPheSerAlaAspValThrPro 200
QY 865 GTCAATTTAATAATATGCNAGAGACCGTTTGGAAATGCTGATGAACCTTCAAGTAGAATT 924
Db 201 ValAsnTyrLysTyrAlaArgThrValTrpAsnGlyAspGluProSerSerArgIle 220
QY 925 CTTGCAACACAGATAGTATCACAGATGTTTCTGGATTATAGTTTGTAGCTGGACAAAC 984
Db 221 LeuAlaAsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsn 240
QY 985 ACGAAGTACCAATTTAGTTTGTAGCAACTATGCTCCATCAACTGGTTTATTTATTTCCCT 1044
Db 241 ThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPhePro 260
QY 1045 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATGATACATAAATAAATAAT 1104
Db 261 TyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGlnTyrLysLeuAsnAsn 280
QY 1105 GGAATGTTCAACAGTTGAGTTGCCACTTCCACTAGTCAAAATAACTACAGCTAAT 1164
Db 281 GlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaGlnGlnThrThrAlaAsn 300
QY 1165 CCAACTCCAGCAGTTGATGAGATTAAAGTTTCTAAATCGTTTATCAGGTTTAAAGATT 1224
Db 301 ProThrProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPhe 320
QY 1225 GCCCAAAACACAAATCGAATTAAGTTTCCCAACGGGTGAAGAAATATGATAAAGTTCCG 1284
Db 321 GlyGlnAsnThrIleGluLeuSerValProThrGlyGluGlyAsnMetAsnLysValAla 340
QY 1285 CCAATGATTGGCAACATTTATCTTAGCTCAAAATGAAAATAATGCTGATAAG 1335
Db 341 PrometIleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLys 357

RESULT 6

US-09-791-537-105581

; Sequence 105581, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 105581

; LENGTH: 645

; TYPE: PRT

; ORGANISM: Mycoplasma gallisepticum

US-09-791-537-105581

Alignment Scores: 1.45e-58 Length: 645

Pred. No.: 1.45e-58 Length: 645


```
QY 388 TTAGAACAACTAAATAATGGCTAAATACTAATTTAGTAATCAGCCATCAACCAAGCTAATACG 447
Db 115 ThrGlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThr 134
QY 448 GATAAAGCACTTTTGTATTAATGAACACCAATTTAGTTGAAGCATACAAAGCACTAAAA 507
Db 135 SerLysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyLysGluLeuLys 154
QY 508 ACCACTTTAGAACACGCTGCTACTAACCTTGAAGGTTTGTATCACTCAACTGCTTATAATCAA 567
Db 155 ThrThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGly 174
QY 568 ATTCGCAATAATTTAGTGGATCTATACATAAAGCTAGCTACTTAAATACTAAACACACTA 627
Db 175 IleLysMetHisLeuSerGlyLeuThrAspAlaGlyLysAlaIleThrThrLysThrLys 194
QY 628 GATCCTCAATATGGGGCAACGCTTTTATGATTTCTAATGATTTACTACAGCTAATAAGAAT 687
Db 195 GluProValGluGlyAspPro--LeuThrAlaGlyAlaValThrMetAlaAsnThrLys 213
QY 688 ATTAATAATACGTTA-----TCACTATTAAATGAACAAAGACTAATGCTGATGCACTA 741
Db 214 IleValGluAlaIleLysAspGluValLeuAsnProLysLysGluAsnAlaThrLysLeu 233
QY 742 TCTAATAGCTTTTATTAATAAAGTGTTCATAATATGAACAAAGTTTGTAGGGACCTTTT 801
Db 234 AlaAspSerPheValLysGlnValLeuValLysGluLysIleThrGlyValGluGluAla 253
QY 802 ACAACGCTAATGTTCAACCTTCAACCTACAGTGTGTTGCTTTTATGCTGATGCTAACA 861
Db 254 HisAsn---LysAlaGlnProAlaAsnTyrSerPheValGlyTyrSerValAspIleThr 272
QY 862 CCCGTC-----AATTATAATATGCAAGAAGGACCGCTTTGG 897
Db 273 GlyThrAlaAsnGlyGlnThrSerIleProAsnTrpAsnTyrAlaGlnArgThrIlePhe 292
QY 898 ---AATGCTGATGACCTTCAAGTAGAATTTCTCAACACG----- 936
Db 293 ThrAsnGlyAspGluPro-----ArgSerValSerAsnThrProValAspGlyGlnThr 310
QY 937 -----AATAGTATCACAGATGTTTCTTGGATTTATATAGTTTGTAGCTGGAACAAACAG 990
Db 311 MetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThrGlyAlaLys 330
QY 991 TACCAATTTAGTTTACCACTAAGTCCATCACTGGTTATTTATATTTCCCTTATAAG 1050
Db 331 TyrThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 350
QY 1051 TTGGTTAAGCAGCTGATGCTTAATAACCTTGTGATTTACATAATTAATAATGGAAT 1110
Db 351 LeuValAsnThrSerAspGlnMetLysLeuGlyLeuGluTyrLysLeuAsnAspAla--- 369
QY 1111 GTTCAACAAGTTGAGTTTGGCACTTCAACTAGTGCATAATACTACAGCTAAT---CCA 1167
Db 370 -----ThrGluProSerAlaIleThrPheGlyAsnGluGlnThrMetAsnGlyLys 386
QY 1168 ACTCAGCAGTGTAGATTAAGTTGCTTAAATCGTTTATCAGTTTAAAGATTGGC 1227
Db 387 ThrProThrValAsnAspIleAsnValAlaLysValThrLeuAlaAsnLeuIlePheGly 406
QY 1228 CAACAACAACTGCAATTAAGTTGTTCCACAGGCTGAGGAATATGAATAAAGTTGCGCCA 1287
Db 407 SerAsnLysIleGluPheSerValProAla-----GluLysValSerPro 421
QY 1288 ATGATTTGCAACATTTATCTTAGCTCAAAATGAATAATGCTGATAAGATC 1338
Db 422 MetIleGlyAsnMetTyrLeuSerSerProAsnAsnTrpAsnLysIle 438
```

RESULT 8

US-08-014-631-7

; Sequence 7, Application US/08014631

; GENERAL INFORMATION:

; APPLICANT: Browning, Glenn F.

```
; TITLE OF INVENTION: MYCOPLASMA PROBES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,631
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 647 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-014-631-7
```

```
Alignment Scores:
Pred. No.: 2,11e-55 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 4 Gaps: 13
```

US-09-147-052-1 (1-1371) x US-08-014-631-7 (1-647)

```
QY 157 GTGGGTTCAACGTCATCGCTCTAGAAATTCGCTGTCTATCTATTACTATAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProAsnProThr 34
QY 211 -----GATGCAACCAATAATATGCG 231
Db 35 ProAsnProAsnProProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
QY 232 CAA-----ACCAATTTAGAACGAGCGCGAATGGAGTTAACAGAT 270
Db 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaArgMetGlyLeuThrThr 74
QY 271 CTAATCAATGTAAAGCGATGACATTTAGCTTACACTACAGACTATGCCAAGATTGAAGCT 330
Db 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGlnAsn 94
QY 331 AGTTTATCATCTGCTTATAGTAGAAGCTGAAACAGTTAAACAATAAATCTTAATGCAACATTA 390
Db 95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThr 114
QY 391 GAACAACATAAATGGCTTAAACTAATTTAGTAATCAGCCATCAACCAAGCTAATACGAT 450
Db 115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY 451 AAAACGACTTTTGTATAATGAACCCAAATTTAGTTGAGTACACAAAGCAGCACTAAAAACC 510
Db 135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyLysGluLeuLysThr 154
QY 511 ACTTTAGAACACGCTGCTACTAAACCTTGAAGGTTTGTATCACTCAACTGCTTATAATCAAAAT 570
```

```
Db 155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY 571 CGCAATAATTTAGTGATCTATACATAAAGCTAGTCTTTAATAACTAAACACTAGAT 630
Db 175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY 631 CCACATAATGGGGAACGCTTTAGATCTCTAATAGATCTACTACAGCTAATAAGAATTT 690
Db 195 ProValGluGlyAspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY 691 AATAATAGTTA-----TCAACTATTAAATCAACAAAGACTAAT-----GCT 732
Db 214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY 733 GATCATATCTAATAGTTTATTAAAAAGTGAATCAAAATAATGAACAAAGTTTGTGA 792
Db 234 AspSerLeuLeuSerSerIleValLysLysIleThrGlyValGluGlu----- 249
QY 793 GGGACTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTTGTGCTTTTAGTGCT 852
Db 250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY 853 GATGTAACACCCGTC-----AATTATAAATATGCAAGAAG 888
Db 267 TrpTyrThrGluLeuLeuLeuAspLysGlnValPheProAsnTyrPaspTyrAlaGlnArg 286
QY 889 ACCGTTTGG---AATGCTGATGAACCTTCAAGTAGAATCTTGTCAACACAG----- 936
Db 287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY 937 -----AATAGTATCACAGATGTTCTTGATTTATGATTTAGTCTGGNACA 981
Db 305 GlycInThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY 982 AACAGGAAGTACCAATTTAGTTTACCAACTATGTCATCAACTGCTTATTTATTTATTC 1041
Db 325 GlyAlaLysTyrThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY 1042 CCTTATAGTTGTTAAAGCAGCTGATGCTTAATACGTTGGATGATACAAATTAAT 1101
Db 345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGlyLeuGluTyrLysLeuAsn 364
QY 1102 AATGGAAATGTTCAACAAGTTGAGTTGCCACTCAACTAGTAGTCA----- 1146
Db 365 Asp-----AlaThrLysProSerAlaIleThrPheGlySer 376
QY 1147 AATAATACTACACTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db 377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY 1207 TTATCAGGTTTAAAGATTTGGCCAAAACACAAATCGAATTAAGTCTTCCACGGGTGAAGA 1266
Db 397 LeuAlaAsnLeuAsnPheGlySerAsnLysIleGluPheSerValProAla----- 413
QY 1267 AATATGAATTAAGTTCGCCAATGATGTCACAACTTATCTTAGCTCAATGAAATAAT 1326
Db 414 -----GluLysValSerProMetIleGlyAsnMetTyrLeuSerSerProAsnAsn 431
QY 1327 GCTGATAAGATC 1338
Db 432 TrpAsnLysIle 435
```

RESULT 9

```
US-08-230-312-7
; Sequence 7, Application US/08230312
; GENERAL INFORMATION:
; APPLICANT: Browning, Glenn F.
; TITLE OF INVENTION: LIVE VACCINE VECTORS BASED ON MYCOPLASMA
; NUMBER OF INVENTIONS: 17
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
```

```
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,312
FILING DATE: April 20, 1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 88012
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 647 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-230-312-7
```

```
Alignment Scores:
Pred. No.: 2,115-55 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 13 Gaps: 13
```

US-09-147-052-1 (1-1371) x US-08-230-312-7 (1-647)

```
QY 157 GTGGGTTCAACGCTGATCGTCTAGAAATTCGGCTGCTATCTATCTACTATAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaAlaSerCysThrThrProThrProAsnProThr 34
QY 211 -----GATGCAAAACCCAAATAATGCG 231
Db 35 ProAsnProAsnProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyaspGly 54
QY 232 CAA-----ACCCAAATTAGAACGACGCGGAATGAGTAAACAGAT 270
Db 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74
QY 271 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAGCT 330
Db 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysThrGlnAsn 94
QY 331 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAACAATTAACCTTAATCAACATTA 390
Db 95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThr 114
QY 391 GAACAACTAAATGCTAAACATAATTTAGATCAGCCATCAACCAAGCTAATACGAT 450
Db 115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY 451 AAACGACTTTTGTATTAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAATAAAC 510
Db 135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY 511 ACTTTAGAACACAGCTGCTACTAACCTTGAAGGTTTGTCACTCACTGCTTATATAATAATT 570
Db 155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY 571 CGCAATAATTTAGTGATCTATACATAAAGCTAGTCTTTAATAACTAAACACTAGAT 630
```



```
Db 175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrLysThrLeuGlu 194
QY 631 CCACATAAATGGGGAACCTTTTACAGATTAATAGAGTACTACAGCTAATAAGAATAAT 690
Db 195 ProValGluGlyAspPro--LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY 691 AATAATACGTTA-----TCAACTATTAAATGAACAAAGACTAAT-----GCT 732
Db 214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY 733 GATGCATTATCTAATAGTTTATTAATAAAGTATTCAAAATATGAACAAAGTTTGTGA 792
Db 234 AspSerLeuLeuSerIleValLysIleThrGlyValGluGlu----- 249
QY 793 GGGACTTTTACAAACGCTTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTAGTGT 852
Db 250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY 853 GATGTAACACCGGTC-----AATTATAAATATGAACAAAGG 888
Db 267 TrpTyrThrGluLeuLeuAspLysGlnValPheProAsnTrpAspTyrAlaGlnArg 286
QY 889 ACCGTTTGG--AATGGTGATGAACCTTCAAGTAGAATCTTGCAACACG----- 936
Db 287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY 937 -----AATAGTATCACAGATGTTCTTGGATTATAGTTTACGTGGAACA 981
Db 305 GlyGlnThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY 982 AACAGAGTAGTGGTTAGTTTATAGCAACTATGTCCTCATCACTGCTGTTATTTATATTC 1041
Db 325 GlyAlaLysTyrThrLeuGluPheThrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY 1042 CCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGATGATACATACAAATTAAT 1101
Db 345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGlyLeuGluTyrLysLeuAsn 364
QY 1102 AATGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCA----- 1146
Db 365 Asp-----AlaThrLysProSerAlaIleThrPheGlySer 376
QY 1147 AATAATACATACAGTATCCAACTCCAGCTTGATGAGATTAAGTTGCTAAATCGTT 1206
Db 377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY 1207 TTATCAGGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTTCCCAACGGTGAAGGA 1266
Db 397 LeuAlaAsnLeuAsnPheGlySerAsnLysIleGluPheSerValProAla----- 413
QY 1267 AATATGAATAAAGTTGCCCAATGATGGCAACATTTATCTAGCTCAAAATGAATAAT 1326
Db 414 -----GluLysValSerProMetIleGlyAsnMetTyrLeuSerSerProAsnAsn 431
QY 1327 GCTGATAGATC 1338
Db 432 TrpAsnLysIle 435

RESULT 10
US-09-791-537-7385
; Sequence 7385, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 7385
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
US-09-791-537-7385

Alignment Scores:
Pred. No.: 2,11e-55 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 22 Gaps: 13

US-09-147-052-1 (1-1371) x US-09-791-537-7385 (1-647)
QY 157 GTGGTTCAACCGTATCGCTAGAAATTCGGCTGCTATGCTATTACTATAAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaIleAlaSerCysThrThrProThrProAsnProThr 34
QY 211 -----GATGCAAAACCCAAATAATAGGC 231
Db 35 ProAsnProAsnProProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
QY 232 CAA-----ACCCAATTAGAGCAGCGCGAATGGAGTTAACAGAT 270
Db 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74
QY 271 CTAATCAATGCTAAAGCGATGACATTAGCTTCCATACAGACTATGCCAAGATTCAAGCT 330
Db 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGlnAsn 94
QY 331 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAACAATAACCTTAATGCAACATTA 390
Db 95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThr 114
QY 391 GAACAACATAAATGGCTAAACATAATTTAGATCAGCCATCAACCAAGCTTAATAGGAT 450
Db 115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY 451 AAAACGACTTTGATAATGAACACCCAAATTTAGTTGAGCATACAAAGCAGCACTAAACACC 510
Db 135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY 511 ACTTTAGAACACCGCTGCTACTAACTTGAAGGTTTGTGCTCACTCACTGCTTATAATCAAA 570
Db 155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY 571 CGCAATAATTTAGTGGATCTATACATAAAGCTAGTAGTTTAACTAACTAAACACACTAGAT 630
Db 175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY 631 CCACATAAATGGGGAACGCTTTTACAGTTTCAATGAGATTACTACAGCTAATAAGAATAAT 690
Db 195 ProValGluGlyAspPro--LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY 691 AATAATACGTTA-----TCAACTATTAAATGAACAAAGACTAAT-----GCT 732
Db 214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY 733 GATGCATTATCTAATAGTTTATTAATAAAGTATTCAAAATATGAACAAAGTTTGTGA 792
Db 234 AspSerLeuLeuSerIleValLysIleThrGlyValGluGlu----- 249
QY 793 GGGACTTTTACAAACGCTTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTAGTGT 852
Db 250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY 853 GATGTAACACCGGTC-----AATTATAAATATGAACAAAGG 888
Db 267 TrpTyrThrGluLeuLeuAspLysGlnValPheProAsnTrpAspTyrAlaGlnArg 286
QY 889 ACCGTTTGG--AATGGTGATGAACCTTCAAGTAGAATCTTGCAACACG----- 936
```

```

Db      287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY      937 -----AATAGTATCACAGATGTTCTTTGGATTTATAGTTTAGCTGGAACA 981
Db      305 GlyGlnThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY      982 AACACGAAGTACCAATTAGTTTGTAGCAACTATGCTCATCACTGGTTATTTATATTC 1041
Db      325 GlyAlaLysTyrThrLeuGluPheThrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY      1042 CCTTATAAGTTGGTTAAAGCAGCTGATGCTAAATAACGTTGGATTACAAATCAATAAAT 1101
Db      345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGluTyrLysLeuAsn 364
QY      1102 AATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGA----- 1146
Db      365 Asp -----AlaThrLysProSerAlaIleThrPheGlySer 376
QY      1147 AATAATATACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db      377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY      1207 TTATCAGGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTGTTCCACGGGTGAAGGA 1266
Db      397 LeuAlaAsnLeuAsnPheGlySerAsnLysIleGluPheSerValProAla----- 413
QY      1267 AATATGAATAAGTTCGCCCAATGATGGCAACATTTATCTTACGCTCAATGAAATAAT 1326
Db      414 -----GluLysValSerProMetIleGlyAsnMetTyrLeuSerSerProAsnAsn 431
QY      1327 GCTGATTAAGATC 1338
Db      432 TrpAsnLysIle 435

RESULT 11
US-09-791-537-34555
; Sequence 34555, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34555
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
US-09-791-537-34555

Alignment Scores:
Pred. No.: 2,11e-55 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 22 Gaps: 13

US-09-147-052-1 (1-1371) x US-09-791-537-34555 (1-647)

QY      157 GTGGGTTCAACGCTGATCGCTAGAAATCGGCTGTATGCTTACTATAAATA----- 210
Db      15 IieGlySerPheValMetIleuAlaAlaSerCysThrProThrProAsnProThr 34
QY      211 -----GATGCAACCCCAATAATGSC 231
Db      35 ProAsnProAsnProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
```

```

QY      232 CAA-----ACCAATATTAGACGACGCGGAATGGAGTTAACAGAT 270
Db      55 GlnglyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74
QY      271 CTAATCAATCCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCTCAAGATTTGAAGCT 330
Db      75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysThrGlnAsn 94
QY      331 AGTTTATCACTCTTATCTAGCTGAAGCTGAACAGTTAAACAATACCTTAATGCAACATTA 390
Db      95 ThrLeuThrLysAlaTyrAspAlaAlaLysThrValLeuAspAsnSerSerSerThrThr 114
QY      391 GAACAACATAAAATGGCTAAACTAATTTAGAAATCAGCCATCAACCAAGCTAATACGGAT 450
Db      115 GlnAsnLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaAlaThrSer 134
QY      451 AAAACAGACTTTTGCATTAATGAACCCCAATTTAGTTGAAGCATPACAAACACCTATAAAC 510
Db      135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY      511 ACTTTAGAACCAACGCTGCTACTAACCTTGAAGTTTCTCATCACTGCTTATATCAATTA 570
Db      155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY      571 CGCAATAATTTAGTGGATCTATACATAAAGCTAGTAGTTTATAACTAAACACACTAGAT 630
Db      175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY      631 CCACTAAATGGGGAACGCTTTTAGATTCTTAATGAGATTACTACAGCTAATAGAATAT 690
Db      195 ProValGluGlyAspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY      691 AATAATACGTTA-----TCAACTATTATTAAGCAAAAGACTAAT-----GCT 732
Db      214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY      733 GATGCAATTATCTAATAGTTTTTATTAATAAAGTATTCAAAATTAATGAACAAAGTTTGT 792
Db      234 AspSerLeuSerSerIleValLysLysIleThrGlyValGluGlu----- 249
QY      793 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTCTTTTAGTCT 852
Db      250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY      853 GATGTAACACCCGTC-----AATTATAATATCAAGAAGG 888
Db      267 TrpTyrThrGluLeuLeuAspLysGlnValPheProAsnTrpAspTyrAlaGlnArg 286
QY      889 ACCGTTTGG---AATGGTGATGAACCTTCAAGTAGAATCTTGGCAACACG----- 936
Db      287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY      937 -----AATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACA 981
Db      305 GlyGlnThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY      982 AACACGAAGTACCAATTTAGTTTGTAGCAACTATGCTCATCACTGGTTATTTATATTC 1041
Db      325 GlyAlaLysTyrThrLeuGluPheThrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY      1042 CCTTATAAGTTGGTTAAAGCAGCTGCTAATACGTTGGATTACAAATCAATAAAT 1101
Db      345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGluTyrLysLeuAsn 364
QY      1102 AATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGA----- 1146
Db      365 Asp -----AlaThrLysProSerAlaIleThrPheGlySer 376
QY      1147 AATAATATACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db      377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY      1207 TTATCAGGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTGTTCCACGGGTGAAGGA 1266
```

```
Db 397 LeuAlaAsnLeuAsnPhcGlySerAsnLysIleGluPheSerValProAla----- 413
QY 1267 AATATGAATAAGTTGCCCAATGATGGCAACATTTATCTTAGCTCAAAATGAAATATAT 1326
Db 414 -----GluLysValSerProMetIleGlyAsnMetYrLeuSerSerProAsnAsn 431
QY 1327 GCTGATGATATC 1338
Db 432 TrpAsnLysIle 435

RESULT 12
US-10-131-591A-5
; Sequence 5, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: MDVgB signal
US-10-131-591A-5

Alignment Scores:
Pred. No.: 7,35e-20 Length: 62
Score: 320.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.76% Indels: 0
DB: 27 Gaps: 0

US-09-147-052-1 (1-1371) x US-10-131-591A-5 (1-62)
QY 1 ATGCACTATTTAGCGGAATGTCATATTTTCTTAGTATCTATATGGTACGAAAC 60
Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuYrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTTGCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCTCCGCCACGAGTGGTTCACCGTATCGTCTA 180
Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAATTC 186
Db 61 Gluphe 62

RESULT 13
US-07-722-860-13
; Sequence 13, Application US/07722860
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 North Washington Street
; CITY: Falls Church
; STATE: Virginia
```

```
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/722,860
; FILING DATE: 19910628
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-101P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-722-860-13

Alignment Scores:
Pred. No.: 4,43e-19 Length: 865
Score: 314.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.50% Indels: 0
DB: 3 Gaps: 0

US-09-147-052-1 (1-1371) x US-07-722-860-13 (1-865)
QY 1 ATGCACTATTTAGCGGAATGTCATATTTTCTTAGTATCTATATGGTACGAAAC 60
Db 1 MethIstYrPheArgArgAsnCysIlePhePheLeuIleValIleLeuYrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTTGCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCTCCGCCACGAGTGGTTCACCGTATCGTCTA 180
Db 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAA 183
Db 61 Glu 61

RESULT 14
US-08-213-449A-13
; Sequence 13, Application US/08213449A
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213,449A
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-108P
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-213-449A-13

Alignment Scores:
Pred. No.: 4.43e-19 Length: 865
Score: 314.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.50% Indels: 0
DB: 6 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-213-449A-13 (1-865)
QY 1 ATGCACATATTTAGCGGAATTCATATTTTCTTATAGTATTCTATATGTCAGCAAC 60
Db 1 MethistyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCCCACCAGTGGTTTCAACCGTGATCCGCTA 180
Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAA 183
Db 61 Glu 61

RESULT 15
US-08-213-449B-13
; Sequence 13, Application US/08213449B
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryobei
; APPLICANT: LI, Y1
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213,449B
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-108P
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-213-449B-13

Alignment Scores:
Pred. No.: 4.43e-19 Length: 865
Score: 314.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.50% Indels: 0
DB: 6 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-213-449B-13 (1-865)
QY 1 ATGCACATATTTAGCGGAATTCATATTTTCTTATAGTATTCTATATGTCAGCAAC 60
Db 1 MethistyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCCCACCAGTGGTTTCAACCGTGATCCGCTA 180
Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAA 183
Db 61 Glu 61

Search completed: October 8, 2003, 19:14:06
Job time: 433.851 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 18:12:37 ; Search time 75.1023 Seconds
(without alignments)
5882.831 Million cell updates/sec

Title: US-09-147-052-1

Perfect score: 2326

Sequence: 1 atgcacatttttagcgga.....gaccgggtacatttttataa 1371

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 1201306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09147052/runat_08102003_154344_29960/app_query.fasta_1.5980
-DB=Published Applications_AA -QFMT=fastan SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09147052.ecgn_1_1_355_@runat_08102003_154344_29960
-NCPU=6 -ICPU=3 -NO_MMAP -LARGUEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FPGAPOP=6 -FPGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	2324	99.9	456	9	US-09-147-052-2	Sequence 2, Appli
2	2317	99.6	456	12	US-09-901-572A-3	Sequence 3, Appli
3	2015	86.6	1086	9	US-09-147-052-4	Sequence 4, Appli
4	1929.5	83.0	384	15	US-10-131-591A-12	Sequence 12, Appli
5	1777	76.4	357	15	US-10-131-591A-23	Sequence 23, Appli
6	320	13.8	62	15	US-10-131-591A-5	Sequence 5, Appli
7	308	13.2	62	15	US-10-131-591A-6	Sequence 6, Appli
8	176.5	7.6	6281	9	US-09-815-242-12996	Sequence 12996, A
9	174	7.5	2086	9	US-09-815-242-5639	Sequence 5639, Ap
10	174	7.5	5795	9	US-09-815-242-12610	Sequence 12610, A
11	173.5	7.5	496	12	US-10-311-879-29	Sequence 29, Appli
12	173.5	7.5	936	8	US-08-781-986A-5249	Sequence 5249, Ap
13	170	7.3	520	15	US-10-056-052-2	Sequence 2, Appli
14	166.5	7.2	2659	12	US-10-311-879-28	Sequence 28, Appli
15	163	7.0	1029	9	US-09-815-242-5885	Sequence 5885, Ap
16	163	7.0	1048	9	US-09-815-242-13083	Sequence 13083, A
17	158	6.8	2434	9	US-09-815-242-5835	Sequence 5835, Ap
18	156	6.7	807	11	US-09-820-843A-108	Sequence 108, App
19	155.5	6.7	2478	9	US-09-815-242-5816	Sequence 5816, Ap
20	155.5	6.7	2478	9	US-09-815-242-12967	Sequence 12967, A
21	149.5	6.4	837	9	US-09-815-242-5883	Sequence 5883, Ap
22	149.5	6.4	875	9	US-09-815-242-13080	Sequence 13080, A
23	149.5	6.4	3158	9	US-09-815-242-12611	Sequence 12611, A
24	148.5	6.4	857	12	US-10-032-585-7857	Sequence 7857, Ap
25	145.5	6.3	2437	9	US-09-815-242-5834	Sequence 5834, Ap
26	145	6.2	1215	9	US-09-815-242-5908	Sequence 5908, Ap
27	145	6.2	1269	9	US-09-815-242-13113	Sequence 13113, A
28	144.5	6.2	596	11	US-09-946-374-243	Sequence 243, App
29	144.5	6.2	596	12	US-10-015-387A-243	Sequence 243, App
30	144.5	6.2	596	12	US-10-063-735-100	Sequence 100, App
31	144.5	6.2	596	12	US-10-006-130A-243	Sequence 243, App
32	144.5	6.2	596	12	US-10-199-672-310	Sequence 310, App
33	144.5	6.2	596	12	US-10-006-172A-243	Sequence 243, App
34	144.5	6.2	596	12	US-10-187-749-310	Sequence 310, App
35	144.5	6.2	596	12	US-10-194-457-310	Sequence 310, App
36	144.5	6.2	596	12	US-10-184-542-310	Sequence 310, App
37	144.5	6.2	596	12	US-10-196-747-310	Sequence 310, App
38	144.5	6.2	596	12	US-10-015-392A-243	Sequence 243, App
39	144.5	6.2	596	12	US-10-017-253A-243	Sequence 243, App
40	144.5	6.2	596	12	US-10-173-689-310	Sequence 310, App
41	144.5	6.2	596	12	US-10-173-690-310	Sequence 310, App
42	144.5	6.2	596	12	US-10-173-691-310	Sequence 310, App
43	144.5	6.2	596	12	US-10-173-692-310	Sequence 310, App
44	144.5	6.2	596	12	US-10-173-694-310	Sequence 310, App
45	144.5	6.2	596	12	US-10-173-698-310	Sequence 310, App

ALIGNMENTS

RESULT 1

US-09-147-052-2
; Sequence 2, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, AND ITS USE
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid

US-09-147-052-2

Alignment Scores:
 Pred. No.: 4,81e-180 Length: 456
 Score: 2324.00 Matches: 456
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 9 Gaps: 0

US-09-147-052-1 (1-1371) x US-09-147-052-2 (1-456)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTCACGAAC 60
 DB 1 Meth1stYrPheArgAsnGlyPhePheLeuLeuValLeuLeuGlyThrAsn 20
 QY 61 TCATCTCGGAGTACCCAAATGTCATCAAGAGAAGTTGTTCCGACGCTCCAGTTGCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCACCGGTGATCCGCTA 180
 DB 41 GluGluGluSerThrPheThrLeuLeuValValSerValGlySerThrValLeuArgLeu 60
 QY 181 GAATTCGGCTGTATGCTTATTAATAAAGATGCAAAACCAATATGATGCAAAACCAA 240
 DB 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
 QY 241 TTGAAGCAGCGCAATGGAGTTAACAGATCTATCAATGCTAAAGCGATGACATAGCT 300
 DB 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
 QY 301 TCACTACAGCATGCTGCAAGATGAGAGTGTATCATCTCTGTTATAGTGAAGCTGAA 360
 DB 101 SerLeuGlnAspThrAlaLysIleGluAlaSerLeuSerSerAlaThrSerGluAlaGlu 120
 QY 361 ACAGTTAACCAATTAACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 420
 DB 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
 QY 421 GAATCAGCCATCAACCAAGTATACGATTAACCAAGTATACCAAGTATACCAAGTAT 480
 DB 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrPheAspAsnGluHisProAsn 160
 QY 481 TTAGTTGAGCATACAAAGCACTTAAACCACTTTAGAACACAGCTGCTACTAACCTGAA 540
 DB 161 LeuValGluAlaThrLysAlaLeuLysThrLeuGluGlnArgAlaThrAsnLeuGlu 180
 QY 541 GPTTCTGCTCACTGCTTATCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 600
 DB 181 GlyLeuSerSerThrAlaThrAsnGlnIleArgAsnAsnLeuValAspLeuThrAsnLys 200
 QY 601 GCTAGTAGTTTAACTAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCT 560
 DB 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
 QY 661 ATGAGATTACTACAGCTAATGAATATTAATATACGTTATCACTATTAATGAACAA 720
 DB 221 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
 QY 721 AAGACTAATGCTGATGCTATCTAATAGTTTATTAATAAGTGAATCAAAATGAATGAA 780
 DB 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 260
 QY 781 CAAAGTTTGTAGGACCTTTTACAAACGCTAATGTTCAACCTTCAACCTTCAACCTT 840
 DB 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnThrSerPheVal 280
 QY 841 GCTTTTAGTGTGATGATACCCGCTCAATATAATATGAAGAAGGACCGTTTGGAT 900
 DB 281 AlaPheSerAlaAspValThrProValAsnThrLysThrAlaArgThrValThrAsn 300
 QY 901 GGTGATGAACCTTCAAGTAGAATCTTCCAAACAGGAATAGTATACAGATGTTCTTGG 960

Db 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 320
 QY 961 ATTTATAGTTTACCTGCAACAAACACAGTACCAATTTAGTTTACCACTATGTCCTCA 1020
 DB 321 IleThrSerLeuAlaGlyThrAsnThrLysThrGlnPheSerPheSerAsnThrGlyPro 340
 QY 1021 TCAACTGGTTTATTTATTTCCCTTATAAGTTGGTTAAACGACGCTGATGCTAATACGTT 1080
 DB 341 SerThrGlyThrLeuThrPheProThrLysLeuValLysAlaAlaAspAlaAsnVal 360
 QY 1081 GGATTCAATACAAATTAATAATGGAATGTTTCAACAGTTGAGTTGCGACTTCAACT 1140
 DB 361 GlyLeuGlnThrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
 QY 1141 AGTGCAAAATTAATACAGCTAATCCAGTCCAGCAGTTCGATGAGATTAAGTTGCTAAA 1200
 DB 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
 QY 1201 ATCGTTTATCAGGTTTAAAGATTGGCCAAACACAAATCGAATTAAGTTGCTCAACGGGT 1260
 DB 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
 QY 1261 GAAGGAATATGAATTAAGTTGGCCCAATGATGGCAACATTTATCTTAGCTCAAAATGAA 1320
 DB 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleThrLeuSerSerAsnGlu 440
 QY 1321 AATAATGCTGATAAGATCCCGGTACCGTCCGACCGGTACATTTTAA 1368
 DB 441 AsnAsnAlaAspLysIleProGlyThrArgArgProGlyThrPheLeu 456

RESULT 2

US-09-901-572A-3
 ; Sequence 3, Application US/09901572A
 ; Publication No. US20030165534A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NIDPOON ZEON CO., LTD.,
 ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
 ; FILE REFERENCE: J209
 ; CURRENT APPLICATION NUMBER: US/09/901,572A
 ; CURRENT FILING DATE: 2003-03-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: Mycoplasma gallisepticum
 ; FEATURE:
 ; OTHER INFORMATION: TTN-1 portion of pNZ40K-S
 US-09-901-572A-3

Alignment Scores:
 Pred. No.: 1,78e-179 Length: 456
 Score: 2317.00 Matches: 454
 Percent Similarity: 99.78% Conservativity: 1
 Best Local Similarity: 99.56% Mismatches: 1
 Query Match: 99.61% Indels: 0
 DB: 12 Gaps: 0

US-09-147-052-1 (1-1371) x US-09-901-572A-3 (1-456)

QY 1 ATGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTCACGAAC 60
 DB 1 Meth1stYrPheArgAsnGlyPhePheLeuLeuValLeuLeuGlyThrAsn 20
 QY 61 TCATCTCGGAGTACCCAAATGTCATCAAGAGAAGTTGTTCCGACGCTCCAGTTGCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACAGTGGGTTCACCGGTGCTA 180
 DB 41 GluGluGluSerThrPheThrLeuLeuValValSerValGlySerThrValLeuArgLeu 60
 QY 181 GAATTCGGCTGTATGCTTATTAATAAAGATGCAAAACCAATATAGTGAATGAA 240

```
Db 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
QY 241 TTAGAAGCAGCGCGAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCT 300
Db 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
QY 301 TCACATAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAA 360
Db 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
QY 361 ACAGTTAAACAATAACCTTAATGCAACATTAGAACAACTAAATAATGGCTAAACTAATTTA 420
Db 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
QY 421 GAATCAGCGATCAACCAAGCTAATACGGATAAAGCGACTTTTGAATGAACACCCAAAT 480
Db 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
QY 481 TTAGTTGAAGCATCAACAGCACTAAACACCCTTTAGAACACAGCTGCTACTAACCTTGA 540
Db 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
QY 541 GGTTCATCACTCACTGCTTATAATCAATTCGAATAATTTAGTGGATCTATACAAATAA 600
Db 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
QY 601 GCTAGTACTTTAATAACTAAACACTAGATCACTCAATTAATGGGGAACGCTTTTATAGTCT 660
Db 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
QY 661 ATGAGATTACTACAGCTAATAAGAAATATTAATATAGTTATCACTTATCACTAATAAGAAC 720
Db 221 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
QY 721 AAGACTAATGCTGATGCTATCTAATAGTTTATTAATAAAGTGAATCAAAATAATGAA 780
Db 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 260
QY 781 CAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTGT 840
Db 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280
QY 841 GCTTTTAGCTGATGTTAACACCGCTCAATTAATAATATGCAAGAGGACGCTTTGGAA 900
Db 281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTyrAsn 300
QY 901 GGTGATGAACCTTCAAGTAGAATCTTGCACACAGCAATAGTATCACAGATGTTTCTTGG 960
Db 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyr 320
QY 961 ATTTATAGTTAGCTGGGAACAAACACGAGTACCAATTTAGTTTGAACACTATGCTCCA 1020
Db 321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
QY 1021 TCACTGCTGTTATTTATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTT 1080
Db 341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
QY 1081 GGATTACATAACAATAATGAAATGTTCAACAAGTTGAGTTTGGCCACTTCAACT 1140
Db 361 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
QY 1141 AGTCAAAATAATPACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTGTCTAAA 1200
Db 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
QY 1201 ATCGTTTTATCAGGTTTAAGATTGGCCAAACACAACTCGAATTAAGTTGTTCCACGGGT 1260
Db 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
QY 1261 GAAGGAAATATGATAAAGTTGCGCAATGATGCGCAACTTTATCTTAGCTCAATGAA 1320
```

```
Db 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
QY 1321 AATAATGCTGATAAGATCCCGGGTACCGTCCGACCGGTACATTTTAA 1368
Db 441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

RESULT 3
US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, NO. US20010014335A1ofu
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

Alignment Scores:
Pred. No.: 5,95e-155 Length: 1086
Score: 2015.00 Matches: 396
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.63% Indels: 0
Db: 9 Gaps: 0

US-09-147-052-1 (1-1371) x US-09-147-052-4 (1-1086)
```

```
QY 181 GAATTCGCTGCTATCTTACTATAAAGATGCAACCCCAATAATGGCCAAACCCAA 240
Db 691 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 710
QY 241 TTAGAAGCAGCGCGAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCT 300
Db 711 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 730
QY 301 TCACATAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAA 360
Db 731 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 750
QY 361 ACAGTTAAACAATAACCTTAATGCAACATTAGAACAACTAAATAATGGCTAAACTAATTTA 420
Db 751 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 770
QY 421 GAATCAGCGATCAACCAAGCTAATACGGATAAAGCGACTTTTGAATGAACACCCAAAT 480
Db 771 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 790
QY 481 TTAGTTGAAGCATCAACAGCTAATAAACCCTTTAGAACACAGCTGCTACTAACCTTGA 540
Db 791 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 810
QY 541 GGTTCATCACTCACTGCTTATAATCAATTCGAATAATTTAGTGGATCTATACAAATAA 600
Db 811 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 830
QY 601 GCTAGTACTTTAATAACTAAACACTAGATCACTCAATTAATGGGGAACGCTTTTATAGTCT 660
Db 831 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 850
```

```

QY 661 AATGAGATTACTACAGCTAATAAGAAATTAATAATACGTTTATCAACTATTAAATGAACAA 720
DB 851 AsnGluLeuThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluIn 870
QY 721 AAGACTAATGCTGATGCTAATCAATAGTTTATTAATAAGTGAATCAAAATAATGA 780
DB 871 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnGlu 890
QY 781 CAAGTTTGTAGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGT 840
DB 891 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 910
QY 841 GCTTTTGTAGTGTGATGAACACCGCTCAATTAATAATATCAATATCAACAGGACGTTTGAAT 900
DB 911 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn 930
QY 901 GGTGATGAACCTTCAAGTAGAATCTTGCACACAGTAATAGTATCAACAGATGTTTCTGG 960
DB 931 GlyAspGluProSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 950
QY 961 ATTTATAGTTTGTAGGCTGAACACACAGTAAGTACCAATTTAGTTTACCAACTATGTTCCA 1020
DB 951 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 970
QY 1021 TCAACTGCTTATTTATATTCCTTATAAGTTGTTAAAGCAGCTGATGCTAATAAGCTT 1080
DB 971 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 990
QY 1081 GGATTACAATAAATTAATGAAATGTTCAACAAGTTGAGTTGGCCTTCAACT 1140
DB 991 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 1010
QY 1141 AGTGCAATTAATACAGCTAATCCAACTCCAGCAGCTGATGAGATTAAGTTGCTAAA 1200
DB 1011 SerAlaAsnAsnThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 1030
QY 1201 ATCGTTTATCAGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTGTTCAACGGGT 1260
DB 1031 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 1050
QY 1261 GAAGGAATATGAATAAGTTGCGCAATGATTGGCAACATTTATCTTACCTCAATGA 1320
DB 1051 GluGlyAsnMetAsnLysValAlaPrometIleGlyAsnIleTyrLeuSerSerAsnGlu 1070
QY 1321 AATAATGCTGATAGATCCCGGGTACCGTCGACCGGTACATTTTA 1368
DB 1071 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 1086

```

RESULT 4

```

US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

```

```

Alignment Scores:
Pred. No.: 4,21e-148 Length: 384
Score: 1929.50 Matches: 382
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.22% Mismatches: 1

```

```

Query Match: 82.95% Indels: 1
DB: 15 Gaps: 1
US-09-147-052-1 (1-1371) x US-10-131-591A-12 (1-384)
QY 181 GAATTCGGCTGTATGCTTACTTAAACCAAGATCAACCAATAATATGCGCAACCA 240
DB 1 GluPhe---CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 19
QY 241 TTAGAAGCAGCGCAATGGAGTTTACAGATCTTAATCAATGCTTAAGCGATGACATTAGCT 300
DB 20 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 39
QY 301 TCATCAACAGCATATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGCTGAAGCTGAA 360
DB 40 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 59
QY 361 ACAGTTTAACAAATTAACCTTAATGCAACATTAGAACAACTAAATAATGCTTAACTAATTTA 420
DB 60 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 79
QY 421 GAATCAGCCATCAACCAAGCTAATACGATTAACAGCTTTTGTATATGACACCCCAAT 480
DB 80 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 99
QY 481 TTAGTTGAACATACAAAGCCTAAACCACTTAAACCACTTTAGAACACGCTGCTACTAACC 540
DB 100 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 119
QY 541 GGTTCGTCACACTGCTTATTAATCAATTCGCAATTAATTTAGTGGATCTATACAATAA 600
DB 120 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 139
QY 601 GCTAGTAGTTTAAATAACACACTAGATCCACTAAATGGGGGAACGCTTTTAGATCTCT 660
DB 140 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 159
QY 661 AATGAGATTACTACAGCTAATGAATTAATAATACGTTTATCAACTATTAAATGAACAA 720
DB 160 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluIn 179
QY 721 AAGACTAATGCTGATGCTAATCAATAGTTTATTAATAAGTGAATCAAAATAATGA 780
DB 180 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnGlu 199
QY 781 CAAAGTTTGTAGGACTTTTACAAACGCTAATGTTTCAACCTTCAAACTACAGTTTGT 840
DB 200 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 219
QY 841 GCTTTTGTAGTGTGATGAACACCGCTCAATTAATAATATGCAAGAGGACCGTTTGAAT 900
DB 220 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn 239
QY 901 GGTGATGAACCTTCAAGTAGAATCTTGCACACACAAATAGTATGATCAACAGTGTCTTCG 960
DB 240 GlyAspGluProSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 259
QY 961 ATTTATAGTTTGTAGGCTGAACACAGTAAGTACCAATTTAGTTTGTAGCAACTATGTTCCA 1020
DB 260 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 279
QY 1021 TCAACTGCTTATTTATATTCCTTATAAGTTGTTAAAGCAGCTGATGCTAATAAGCTT 1080
DB 280 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 299
QY 1081 GGATTACAATAAATTAATGAAATGTTTCAACAGTGTGAGTTTGGTTCCTCAACT 1140
DB 300 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 319
QY 1141 AGTGCAATTAATACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAA 1200
DB 320 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 339

```


QY 1201 ATCGTTTATCAGGTTTAAAGTTTGGCCAAAACACAATCGAATTAAGTGTTCACACGGGT 1260
 Db 340 IIEValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 359
 QY 1261 GAAGCAATATGAATAAAGTTGCGCAATGATGTCGCAACATTTATCTTAGCTCAATGAA 1320
 Db 360 GUGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 379
 QY 1321 AATAATGCTGATAAG 1335
 Db 380 AsnAsnAlaAspLys 384

RESULT 5
 US-10-131-591A-23
 ; Sequence 23, Application US/10131591A
 ; Publication No. US20030059799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nippon Zeon Co., Ltd.,
 ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
 ; FILE REFERENCE: J209
 ; CURRENT APPLICATION NUMBER: US/10/131,591A
 ; CURRENT FILING DATE: 2002-08-15
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 23:
 ; LENGTH: 357
 ; TYPE: PRT
 ; ORGANISM: Mycoplasma gallisepticum
 ; FEATURE:
 ; OTHER INFORMATION: Modified T71-1 portion (downstream of BglI) of
 ; OTHER INFORMATION: PN240R-S
 US-10-131-591A-23

Alignment Scores:
 Pred. No.: 9,38e-136 Length: 357
 Score: 1777.00 Matches: 352
 Percent Similarity: 98.60% Conservative: 0
 Best Local Similarity: 98.60% Mismatches: 5
 Query Match: 76.40% Indels: 0
 DB: 15 Gaps: 0

US-09-147-052-1 (1-1371) x US-10-131-591A-23 (1-357)

QY 265 ACAGATCTAATCAATGCTAAGCGATGACATTAGCTTACATACAGACTATGCCAAGATT 324
 Db 1 ThrAspLeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIle 20
 QY 325 GAAGTAGTTATCATCTGCTTATAGTGAAGCTGAACAGATTAAACAATCACTTAATGCA 384
 Db 21 GluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsnAsnLeuGlnAla 40
 QY 385 ACATTAGAACAACTAAAATGGCTAAAACATAATTTAGAAATCAGCCATCAACCAAGCTAAT 444
 Db 41 ThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsn 60
 QY 445 ACGGATAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAACACTA 504
 Db 61 ThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeu 80
 QY 505 AAAACCACTTAGAACACGCTGCTACTAACCTTGAAGGTTTGTGATCACTGCTTATAAT 564
 Db 81 LysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsn 100
 QY 565 CAATTCGCAATAATTTAGTGGATCTATACATAAAGCTAGTAGTTTAATTAACATAACA 624
 Db 101 GlnIleArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThr 120
 QY 625 CTAGATCCACTAAATGGGGACCGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAG 684
 Db 121 LeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLys 140
 QY 685 ATATATTAATATACGTTTATCACTATTAAATGACAAAAGACTAATGCTGATGCTATTACT 744
 Db 1 ATGCACTATTTTAGCGGAATTCATATTTTCCCTTATAGTATTCTATATGTTACGACAA 60

Db 141 AsnIleGlnAsnThrLeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSer 160
 QY 745 AATAGTTTTTAAAAAGTGTATCAAAATAATGAACAAAAGTTTGTAGGCACTTTTACA 804
 Db 161 AsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThr 180
 QY 805 AACGCTAATGTTCAACCTTCAACACTACAGTTTGTGCTTTTGTGCTGTAGTGTACACCC 864
 Db 181 AsnAlaAsnValGlnProSerGlnTyrSerPheValAlaPheSerAlaAspValThrPro 200
 QY 865 GTCATTTATAATATGCAAGAGGACCGTTTGGATGCTGATGAACCTTCAAGTAGAATT 924
 Db 201 ValAsnTyrLysTyrAlaArgThrValIrpAsnGlyAspGluProSerSerArgIle 220
 QY 925 CTGCAAAACAGAAATAGTATCACAGATGTTTCTTGGATTATAGTTTACGTGGAACAAC 984
 Db 221 LeuAlaAsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsn 240
 QY 985 ACGAAGTACCAATTTAGTTTGTAGCAACTATGCTCATCACTCACTGGTGTATTTATTCCT 1044
 Db 241 ThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPhePro 260
 QY 1045 TATAAGTTGGTTAAAGCAGCTGATCTAATAACGTTGGTTTACATACAAATTAATAT 1104
 Db 261 TyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGlnTyrLysLeuAsnAsn 280
 QY 1105 GGAATGTTCAACAAGTTGAGTTGCCACTTCACTAGTCAAAATAACTACACTAAT 1164
 Db 281 GlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaGlnGlnThrThrAlaAsn 300
 QY 1165 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATTCAGGTTTAAGATT 1224
 Db 301 ProThrProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPhe 320
 QY 1225 GGCCAAAACACAATCGAATTAAGTGTTCACACGGGTGAAGNAATATCAATAAAGTCGG 1284
 Db 321 GlyGlnAsnThrIleGluLeuSerValProThrGlyGluGlyAsnMetAsnLysValAla 340
 QY 1285 CCAATGATGGCAACATTTATCTTAGCTCAAAATGAAATAATGCTGATAAG 1335
 Db 341 PrometIleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLys 357

RESULT 6
 US-10-131-591A-5
 ; Sequence 5, Application US/10131591A
 ; Publication No. US20030059799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nippon Zeon Co., Ltd.,
 ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
 ; FILE REFERENCE: J209
 ; CURRENT APPLICATION NUMBER: US/10/131,591A
 ; CURRENT FILING DATE: 2002-08-15
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5:
 ; LENGTH: 62
 ; TYPE: PRT
 ; ORGANISM: Marek's disease gammaherpesvirus
 ; FEATURE:
 ; OTHER INFORMATION: MDVgB signal
 US-10-131-591A-5

Alignment Scores:
 Pred. No.: 7,31e-18 Length: 62
 Score: 320.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.76% Indels: 0
 DB: 15 Gaps: 0

US-09-147-052-1 (1-1371) x US-10-131-591A-5 (1-62)

QY 1 ATGCACTATTTTAGCGGAATTCATATTTTCCCTTATAGTATTCTATATGTTACGACAA 60

```

Db      1 MethHsTyrPheArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY      61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAGAGTGTTCGAGCGTCCAGTGTGCT 120
Db      21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY      121 GAGGAAGAGTCTACGTTTATCTTGTGTCCTCCACAGTGGGTCAACCGTGATCCGCTGA 180
Db      41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY      181 GAATTC 186
Db      61 GluPhe 62

```

RESULT 7

```

US-10-131-591A-6
; Sequence 6, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: NIPPON ZEON CO., LTD.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified vgb signal
US-10-131-591A-6

```

Alignment Scores:

```

Pred. No.: 6.85e-17 Length: 62
Score: 308.00 Matches: 60
Percent Similarity: 96.77% Conservative: 0
Best Local Similarity: 96.77% Mismatches: 2
Query Match: 13.24% Indels: 0
DB: 15 Gaps: 0

```

US-09-147-052-1 (1-1371) x US-10-131-591A-6 (1-62)

```

QY      1 ATGCACATATTTAGCGGGAATTCATATTTTCCTTATAGTATTCTATATGTCAGCAAC 60
Db      1 MethHsTyrPheArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrGln 20
QY      61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAGAGTGTTCGAGCGTCCAGTGTGCT 120
Db      21 SerSerProSerThrGlnGlnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY      121 GAGGAAGAGTCTACGTTTATCTTGTGTCCTCCACAGTGGGTCAACCGTGATCCGCTGA 180
Db      41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY      181 GAATTC 186
Db      61 GluPhe 62

```

RESULT 8

```

US-09-815-242-12996
; Sequence 12996, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.

```

```

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12996

```

Alignment Scores:

```

Pred. No.: 6.62e-06 Length: 6281
Score: 176.50 Matches: 98
Percent Similarity: 37.36% Conservative: 66
Best Local Similarity: 22.32% Mismatches: 178
Query Match: 7.59% Indels: 97
DB: 9 Gaps: 16

```

US-09-147-052-1 (1-1371) x US-09-815-242-12996 (1-6281)

```

QY      223 AATAATGGCCAAACCAATTTAGAGCGAGCGCAATGGAGTTACACATCTAATCAATGCT 282
Db      266 SerAsnThrSerThrGlnLeuAsnThrAlaMetAlaAsnLeuGlnAsnGlyIleAsnAsp 285
QY      283 AAAGCGATGACATTTAGCTTCTACTACAAGACTATGCCAAGATTGCAAGCTAGTTTATCATCT 342
Db      286 LysThrAsnThrLeuAlaSer---GluAsnTyrHisAspAlaAspSerAspLysLysThr 304
QY      343 GCTTATAGTGAAGCTCAACA---GTTAACAATAACCTTAATCAACA 387
Db      305 AlaTyrThrGlnAlaValThrAsnAlaGluAsnIleLeuAsnLysAsnSerGlySerAsn 324
QY      388 TTAGAACAACATAAAATGGCTAAACTAATTTAGAAATCAGCCATCAACCAAGCTAATACG 447
Db      325 LeuAsp-----LysThrAlaValGluAsnAlaLeuSerGlnValAlaAsn 339
QY      448 GATAAAGCAGCTTTTCATTAATGAACACCCAAATTTAGTGAAGCATCAAAAGCAGCTAAAA 507
Db      340 AlaLysGlyAlaLeuAsnGlyAsnHis---AsnLeuGluGlnAlaLysSerAsnAlaAsn 358
QY      508 ACCACTTTAGAACAAACGCTGCTACTAACCTTGAAGCTTTGTCTCATCACTGCTTATTAATCAA 567
Db      359 ThrThrIle-----AsnGlyLeuGlnHisLeuThrThrAlaGlnLysAspLys 374
QY      568 ATTGCCAATAATTTAGTGGATCTATACAAATAAGCTAGT---AGTTTAAATACTAAACA 624
Db      375 LeuLysGlnValGlnGlnAlaGlnValAlaGlyValAlaAspThrValLysSerSer 394
QY      625 CTAGATCCACTAAATGGG-----GGAACGCTTTTATAGATCTAATGAGTACTACAGCT 678
Db      395 AlaAsnThrLeuAsnGlyAlaMetGlyThrLeuArgAsnSerIleGlnAspAsnThrAla 414
QY      679 AATAAGATATTATAATACGTTTATCACTATTATTAATGAACAAAGACT----- 726
Db      415 ThrLysAsnGlyGlnAsnTyrLeuAspAlaThrGluArgAsnLysThrAsnTyrAsnAsn 434

```

```
QY 727 -----AATGCTGATGCATTA 741
Db 435 AlaValAspSerAlaAsnGlyValIleAsnAlaThrSerAsnProAsnMetAspAlaAsn 454
QY 742 TCTAATATTTTATTAAAAAGTCATCAAAATATGAACAAAGCTTTTGTAGGAGCTTTT 801
Db 455 AlaIleAsnGlnIleAlaThrGlnValThrSerThrIlysAsnAlaLeuAspGlyThrHis 474
QY 802 ACAACGCTAATGTTCACCTTCAACCTACAGTTTGTCTTTAGTGTCTGATGTAACA 861
Db 475 AsnLeuThrGlnAlaLysGlnThr-----AlaThrAsnAlaIleAspGly 489
QY 862 CCCGTCATTAATATATGCAAGAGCAGCCGTTTGGATGGTGATGAACCTTCAAGTAGA 921
Db 490 AlaThrAsnLeuAsnLysAlaGlnLysAspAlaLeuLysAlaGlnValThrSerAlaGln 509
QY 922 ATTCCTTCAACACCAAGTAGATATACAGATGCTTCTTGGATTTTATAGTTAGCT----- 975
Db 510 ArgValAlaAsnValThrSerIleGlnGlnThrAlaAsnGluLeuAsnThrAlaMetGly 529
QY 976 -----GGAACAAACACGAAGTACCAATTTAGTTT 1005
Db 530 GlnLeuGlnHisGlyIleAspAspGluAsnAlaThrLysGlnThrGlnLysTyArgAsp 549
QY 1006 AGCAACTATGTCATCACTGCTTATATATTTCCCTTATAAGTTGGTTAAAGCAGCT 1065
Db 550 AlaGluGlnSerLysThrAlaThr-----AspGlnAlaValAlaAlaAla 565
QY 1066 GATGCT-----AATAACGTTGGATTACAAATACAAA 1095
Db 566 LysAlaIleLeuAsnLysGlnThrGlySerAsnSerAspLysAlaAlaValAspArgAla 585
QY 1096 TTAAAT-----AATGAAATGTTCAACAGTTGAGTTT 1128
Db 586 LeuGlnGlnValThrSerThrLysAspAlaLeuAsnGlyAspAlaLysLeuAlaGluAla 605
QY 1129 GCCACTTCACTAGTGCAT-----AATACTACAGCTAATCCAACT 1170
Db 606 LysAlaAlaAlaLysGlnAsnLeuGlyThrLeuAsnHisIleThrAsnAlaGlnArgThr 625
QY 1171 CCAGCAGTGTATGAGTTAAAGTTGCTAAATCTGTTTATCAGGTTTAAAGATTTGGCAA 1230
Db 626 AspLeuGluGlyGlnIleAsnGlnAlaThrThrVal-----AspGlyVal 640
QY 1231 AACCAATCGAATTAAGTTTCCACGGGTGAAGAAATATGAAT----- 1275
Db 641 AsnThrValLysThrAsnAlaAsnThrLeuAspGlyAlaMetAsnSerLeuGlnGlySer 660
QY 1276 -----AAAGTTGCCCAATGATGGCAACATTTATCTTACTCTCAATGAAAT 1323
Db 661 IleAsnAspLysAspAlaThrLeuArgAsnGlnAsnTyThrLeuAspAlaAspGluSer 679
```

RESULT 9

```
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5639
; LENGTH: 2086
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5639
```

Alignment Scores:

Pred. No.:	8,79e-06	Length:	2086
Score:	174.00	Matches:	104
Percent Similarity:	36.42%	Conservative:	77
Best Local Similarity:	20.93%	Mismatches:	194
Query Match:	7.48%	Indels:	122
DB:	9	Gaps:	19

US-09-147-052-1 (1-1371) x US-09-815-242-5639 (1-2086)

```
QY 49 TATGCTAGCACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGC 108
Db 886 HisAlaLeuAsnAsnLeuThrSerIleAsnAsnAlaGlnLysArgAspLeuThrThrLys 905
QY 109 GTCCAGTGTCTGAGGAAGAGTCTACGTTTATCTTGTCCCCACCAGTGGTTCACACC 168
Db 906 Ile-----AspGlnAlaThr----- 910
QY 169 GTGATCCGCTGATGAATTCGGCTGTGTCTATTACTATAAAGATGCACCAACCAATAAT 228
Db 911 -----ThrValAlaGlyValGluAlaValSerAsnThr 921
QY 229 GCCCAACCCCAATTAGAGCAGCGCAATGGAGTTACAGATCAATCAATGCTAAACCG 288
Db 922 Gly---ThrGlnLeuAsnThrAlaMetAlaAsnLeuGlnAsnGlyIleAsnAspLysAla 940
QY 289 ATGACATTAGCTTCACACAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTAT 348
Db 941 AsnThrLeuAlaSer---GluAsnTyHisAspAlaAspSerAspLysLysThrAlaThr 959
QY 349 AGTGAAGCTGAACAC-----GTTAACATACCTTAATGCAACATAGAA 393
Db 960 ThrGlnAlaValThrAsnAlaGluAsnIleLeuAsnLysAsnSerGlySerAsnLeuAsp 979
QY 394 CAACTAAATGGCTAAACCTAAATTAGAATCAGCCATCAACCAAGCTAATACCGATAAA 453
Db 980 -----LysAlaAlaValGluAsnAlaLeuSerGlnValThrAsnAlaLys 994
QY 454 ACGACTTTTGATAGCAACCCCAATTTAGTTGAAGCATACAAAGCAGCTAAACCACT 513
Db 995 GlyAlaLeuAsnGlyAsnHis---AsnLeuGluGlnAlaLysSerAsnAlaAsnThrThr 1013
QY 514 TTAGAACAACGCTGCTACTTAACCTTTGAAGGTTTGTATCACTCACTGCTTATAATCAATTCGC 573
Db 1014 Ile-----AsnGlyLeuGlnHisLeuThrThrAlaGlnLysAspLysLysLys 1029
QY 574 AATAATTTAGTGGATCTATACAATAAAGCTAGT---AGTTTAATACTAAACACTAGAT 630
Db 1030 GlnGlnValGlnGlnAlaGlnAsnValAlaGlyValAspThrValLysSerSerAlaAsn 1049
QY 631 CCACATAATGGG-----GGAACGCTTTTACATTTCTAATGAGATTACTACAGCTAATAAG 684
Db 1050 ThrLeuAsnGlyAlaMetGlyThrLeuArgAsnSerIleGlnAspAsnThrAlaThrAsn 1069
```



```

QY 631 CCACTAATATGG-----GGAACGCTTTTATAGATCTTAATCAGATTACTACAGCTAATAAG 684
Db 3327 ThrLeuAsnGlyAlaMetGlyThrLeuArgAsnSerIleGlnAspAsnThrAlaThrAsn 3346
QY 685 ATATATTAATAATACGTTATCACTATTAATGAACAAAGACT----- 736
Db 3347 AsnGlyClnAsnThrLeuAspAlaThrGluSerAsnLysThrAsnThrAsnAsnAlaVal 3366
QY 727 -----AATGCTGATGCTATATCTAAT 747
Db 3367 AspSerAlaAsnGlyValIleAsnAlaThrSerAsnProAsnMetAspAlaAsnAlaIle 3386
QY 748 AGTTTTATTAATAAAGTGTCAAAATATGACAAAGTTTGTAGGACTTTTCAAAC 807
Db 3387 AsnGlnIleAlaThrGlnValThrSerThrLysAsnAlaLeuAspGlyThrHisAsnLeu 3406
QY 808 GCTAATGTTCAACCTTCAAACTACAGTTGTTGCTTTAGTCTGCTGATGTAACACCGCTC 867
Db 3407 ThrGlnAlaLysGlnThr-----AlaThrAsnAlaIleAspGlyAlaThr 3421
QY 868 AATTATAATATGCAAGAAGACCGTTTGAATGGTGAACCTTCAAGTAGAATTCCT 927
Db 3422 AsnLeuAsnLysAlaGlnLysAspAlaLeuLysAlaGlnValThrSerAlaGlnArgVal 3441
QY 928 GCAACACAGTAATAGTATCACAGATGTTCTTGATGATTTATAGTTAGCT----- 975
Db 3442 AlaAsnValThrSerIleGlnThrAlaAsnGluLeuAsnThrAlaMetGlyGlnLeu 3461
QY 976 -----GGAACAAACACGGAAGTACCAATTTAGTTTTCAGCAAC 1011
Db 3462 GlnHisGlyIleAspAspGluAsnAlaThrLysGlnThrClnLysThrArgAspAlaGlu 3481
QY 1012 TATGGTCCATCACTGTTATTTATATATCCCTTATAAGTTGGTTAAAGCAGCTGATGCT 1071
Db 3482 GlnSerLysLysThrAlaTyr-----AspGlnAlaValAlaAlaLysAla 3497
QY 1072 -----AATAACGTTGGATTACAAATTAAT 1101
Db 3498 IleLeuAsnLysGlnThrGlySerAsnSerAspLysAlaAlaValAspArgAlaLeuGln 3517
QY 1102 -----AATGGAATGTTCAACAAGTTGAGTTTGGCACT 1134
Db 3518 GlnValThrSerThrLysAspAlaLeuAsnGlyAspAlaLysLeuAlaGluAlaLysAla 3537
QY 1135 TCACTAGTGCAAT-----AATACTACGCTAATCCAACTCCAGCA 1176
Db 3538 AlaAlaLysGlnAsnLeuGlyThrLeuAsnHisIleThrAsnAlaGlnArgThrAlaLeu 3557
QY 1177 GTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGTTTAAAGTTTGCCCAAAACACA 1236
Db 3558 GluGlyGlnIleAsnGlnAlaThrThrVal-----AspGlyValAsnThr 3572
QY 1237 ATCGAATAAGTGTCCAAACGGTGAAGGAAATATGAAT----- 1275
Db 3573 ValLysThrAsnAlaAsnThrLeuAspGlyAlaMetAsnSerLeuGlnGlySerIleAsn 3592
QY 1276 ---AAGTTGCGCAATGATGGCAACATTATCTTAGCTCAATGAAT 1323
Db 3593 AspLysAspAlaThrLeuArgAsnGlnAsnThrLeuAspAlaAspGluSer 3609

```

RESULT 11

```

US-10-311-879-29
; Sequence 29, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: Coxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 29
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-311-879-29

Alignment Scores:
Pred. No.: 7.6e-06 Length: 496
Score: 173.50 Matches: 111
Percent Similarity: 35.90% Conservative: 66
Best Local Similarity: 22.52% Mismatches: 185
Query Match: 7.46% Indels: 131
DB: 12 Gaps: 23

```

US-09-147-052-1 (1-1371) x US-10-311-879-29 (1-496)

```

QY 4 CACATATTTAGGGGAATTGCATATTTTCTCTATAGTTATCTTA----- 48
Db 9 HisAlaIleArgLysLysSerIleGlyValAlaSerValLeuValGlyThrLeuIleGly 28
QY 49 TATGGTACGAACTCATCT-----CCGAGTACCCAAATGTGACATCAAGAGAA 96
Db 29 PheGlyLeuLeuSerSerLysGluAlaAspAlaSerGluAsnSerValThrGlnSerAsp 48
QY 97 GTTGTTCGAGCTCCAGTTGCTGAGGAAGAGCTCTACGTTTATCTTTCTCCCCACCA 156
Db 49 SerAlaSerAsnGluSerLysSerAsnAspSerSerValSerAlaAlaProLysThr 68
QY 157 GTGGGTTCACACCGTATCCGCTAGATTCGGCTGTATGCTATTAATAAAAGATGCA 216
Db 69 AspAspThrAsnVal-----SerAspThrLysThrSerSer 80
QY 217 AACCCAAATATGCGCAACCCCAATTA-----GAAGCAGCGGAATGGAGTTAAACAGAT 270
Db 81 AsnThrAsnAsnGlyGluThrSerValAlaGlnAsnProAlaGlnGlnGluThrThrGln 100
QY 271 CTAATCAATCTAAACGGATGACATTAGCTTCACTACACACATGCCAAGATTGAAGCT 330
Db 101 SerSerSerThrAsnAlaThrThr----- 108
QY 331 AGTTTATCTCTGCTTATAGTAGTGAAGCTGAACACAGTTAAACAATAACCTTAATGCAACATTA 390
Db 109 GluGluThrProValThrGlyGluAlaThrThrThrThrAsnGlnAlaAsnThrPro 128
QY 391 GAAACAATAAATGGCTAAACTAATTTAGATATCAGCCATCAACCAAGCTAATACGGAT 450
Db 129 AlaThrThrGlnSerSerAsnThrAsnAlaGluLeuValAsnGln---ThrSerAsn 147
QY 451 AAACACGACTTTTGTATATGAACACCCCAAAATTTAGTTGAACCATACAAGCAGCTAAACACC 510
Db 148 GluThrThrPheAsnAsp-----ThrAsnThrValSerSerValAsnSer----- 162
QY 511 ACTTTAGAACACAGCTGCTACTAATCTTGAAGTTTGTGCATCAACT----- 555
Db 163 -----ProGlnAsnSerThrAsnAlaGluAsnValSerThrThrGlnAspThrSerThr 180
QY 556 -----GCTTATATCAATTCGCAATAATTTAGTGATCTATACAATAAA--- 600
Db 181 GluAlaThrProSerAsnAsnGluSerAlaProGlnSerThrAspAlaSerAsnLysAsp 200
QY 601 -----GCTAGTAGTTTAAATAACT 618
Db 201 ValValAsnGlnAlaValAsnThrSerAlaProArgMetArgAlaPheSerLeuAlaAla 220
QY 619 AAAACACTAGATCCACTAAATGGGGNAACGCTTTTAGATTCTTAATGAGATTACTACAGCT 678
Db 221 ValAlaAspAlaProAlaAlaGlyThrAspIle---ThrAsnGlnLeuThrAsnVal 239
QY 679 AATAAGAATAATTAATACGTTATCACTATTAATAATGAACAA-----AAGACT 726
Db 240 ThrValGlyIleAspSerGlyThrThrValTyrProHisGlnAlaGlyThrValLysLeu 259
QY 727 AAT---GCTGATGCATATATCTAATAGTTTATTAATA-----AAAGTGAATCA 771

```

```

Db 260 AsnTyrGlyPheSerValProAsnSerAlaValLysGlyAspThrPheLysIleThrVal 279
Qy 772 AATAATGAACAAGATTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTAC 831
Db 280 ProLysGluLeuAsnLeuAsnGlyValThrSerThAlaLysValProPro----- 296
Qy 832 AGTTTGTCTTTTAGTGTGTATGATGACACCGCTCAATTTATAATATGCAAGAGGACC 891
Db 296 ----- 296
Qy 892 GTTTGGAATGATGATGAACCTTCAAGTAGAATTTTGCRAACACGAGTATGATACAGAT 951
Db 297 IleMetAlaGlyAspGln-----ValLeuAlaAsnGlyValIleAspSerAsp 312
Qy 952 GTTTCTTGGATTTATGTTTAGTGTGA---ACAACACGAGAGTACCATAATTTAGTTTAGG 1008
Db 313 GlyAsnValIleThrPheThrAspTyrValAsnThrLysAspValLysAlaThr 332
Qy 1009 AACTATGGTCCCACTGTTTATTTATTTCCCTTATAGTTGTTTAAAGCAGCTGAT 1068
Db 333 LeuThrMetProAla-----TyrIle-----Asp 340
Qy 1069 GCTAATAACGTGGATTACATAACAATAATAATGAATGAAATCTTCAACAAGTTGAGTTT 1128
Db 341 ProGluAsnVal-----LysLysThrGlyAsnVal----- 350
Qy 1129 GCCACTTCAACAGTCCAAATATACTACAGCTAATCCACAGCTCCAGCAGTTGATGAGATT 1188
Db 351 ThrLeuAlaThrGlyIleGlySerThrThrAlaAsnLysThrValLeuValAspTyrGlu 370
Qy 1189 AAGTTGCTTAAATC-----GTTTATCAGTTTAAAGTTTGGC 1227
Db 371 LysTyrGlyPheThrAsnLeuSerIleLysGlyThrIleAspGlnIleAspLysThr 390
Qy 1228 CAACACACATCGAATTAAGTGTCCACGGGTGAAGGAAATATGAATAAGTTGCGCA 1287
Db 391 AsnAsnThrTyrArgGlnThrIleTyrValAsnProSerGlyAspAsnValIleAlaPro 410
Qy 1288 ATGATT---GGCAACATTTATCTTACCTACCTCAATGAAT 1323
Db 411 ValLeuThrGlyAsnLeuLysProAsnThrAspSerAsn 423

```

RESULT 12

```

US-08-781-986A-5249
; Sequence 5249, Application US/08/781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446

```

```

; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5249

```

Alignment Scores:

```

Pred. No.: 8,44e-06 Length: 936
Score: 173.50 Matches: 110
Percent Similarity: 35.29% Conservative: 64
Best Local Similarity: 22.31% Mismatches: 188
Query Match: 7.46% Indels: 131
DB: 8 Gaps: 22

```

US-09-147-052-1 (1-1371) x US-08-781-986A-5249 (1-936)

```

Qy 4 CACTATTTTAGCGGAATGTCATATTTTCTCTTATAGTATTCTA----- 48
Db 18 HisAlaIleArgLysLysSerIleGlyValAlaSerValLeuValGlyThrLeuIleGly 37
Qy 49 TATGTCAGCACTCATCT-----CGAGTACCCAAATGTGACATCAAGAA 96
Db 38 PheGlyLeuLeuSerSerLysGluAlaAspAlaSerGluAsnSerValThrGlnSerAsp 57
Qy 97 GTTGTTCGAGCGTCAGTGTCTGAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCCA 156
Db 58 SerAlaSerAsnGluSerLysSerAsnAspSerSerValSerAlaAlaProLysThr 77
Qy 157 GTGGTTTCAACCGTGATCCGTCCTAGATTCGGCTGTATGCTCTATCTATAAAGATGCA 216
Db 78 AspAspThrAsnVal-----SerAspThrLysThrSerSer 89
Qy 217 AACCCAAATAATGGCCAAACCAATTA-----GAAGCAGCGCAATGGAGTTAACAGAT 270
Db 90 AsnThrAsnAsnGlyGluThrSerValAlaGlnAsnProAlaGlnGlnGluThrGln 109
Qy 271 CTATCAATGCTAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 330
Db 110 SerSerSerThrAsnAlaThr----- 117
Qy 331 AGTTTATCATCTGCTTATAGTGAAGCTGAACAGCTTAACATAACCTTAATGCAACATTA 390
Db 118 GluGluThrProValThrGlyGluAlaThrThrThrThrThrAsnGlnAlaAsnThrPro 137
Qy 391 GAACAACATAAATGGCTAAACTTAATTTAGAAATCAGCCATCAACCAAGCTTAATACCGAT 450
Db 138 AlaThrThrGlnSerSerAsnThrAsnAlaGluGluLeuValAsnGlnThrSerAsnGlu 157
Qy 451 AAAACGAGCTTTGATATATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 510
Db 158 ThrThrSerAsnAspThr-----AsnThrValSerSerValAsnSer----- 171
Qy 511 ACTTTAGAACAACGCTGCTACTTAACCTTGAAGGTTTGTCAATCAACT----- 555
Db 172 -----ProGlnAsnSerThrAsnAlaGluAsnValSerThrThrGlnAspThrSerThr 189
Qy 556 -----GCTTATAATCAATTTGCAATAAATTTAGTGATCTATACATAA--- 600
Db 190 GluAlaThrProSerAsnAsnGluSerAlaProGlnSerThrAspAlaSerAsnLysAsp 209
Qy 601 -----CCTAGTAGTTTAACT 618
Db 210 ValValAsnGlnAlaValAsnThrSerAlaProArgMetArgAlaPheSerLeuAlaAla 229
Qy 619 AAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTTAATGAGATTACTACAGCT 678

```


QY	805	ACGCTAATAGTTTCAACCTTCAAACTACAGTTTGTGCTTTAGTCTGATGTACACCC	864
Db	252	ThrAlaLysValProPro	
QY	865	GTCAATTATAAATATGCAAGAGCGGTTTGGAAATGCGTATGAACCTTCAAGTAGAATT	924
Db	258	-----IleMetaIaGlyAspGln	
QY	925	CTTGGAAACACAGATAGTATCACAGATGTTTCTGGATTTATGATTTAGCTTGA	981
Db	265	LeuAlaAsnGlyValIleAspSerAspGlyAsnValIleTyrThrPheThrAspTyrVal	284
QY	982	AACACGAAGTACCAATTTAGTTTACCACTATGCTCCATCACTGGTATTATATATTC	1041
Db	285	AsnThrLysAspAspValLysAlaThrLeuThrMetProAla	300
QY	1042	CCTTATAAGTTGGTTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAT	1101
Db	301	-----AspProGluAsnVal	
QY	1102	AATGGAAATGTTCAACAAGTTGAGTTTCCCACTTCAACTAGTGCAATATAATACACGCT	1161
Db	308	ThrGlyAsnVal	322
QY	1162	AATCCAATCCAGCAGTTGATGAGATTAAGTTGCTAAATC	1203
Db	323	AsnLysThrValLeuValAspTyrGluLysTyrGlyLysPheTyrAsnLeuSerIleLys	342
QY	1204	--GTTTATCAGGTTTAAGATTTGGCCAAACACATCGAATTAAGTGTTCACACGGGT	1260
Db	343	GlyThrIleAspGlnIleAspLysThrAsnAsnThrTyrArgGlnThrIleTyrValAsn	362
QY	1261	GAAGGAATATGAATAAGTTGCCCAATGATT---GCACACATTTATCTTAGCTCAAT	1317
Db	363	ProSerGlyAspAsnValIleAlaProValLeuThrGlyAsnLeuLysProAsnThrAsp	382
QY	1318	GAATAAT 1323	
Db	383	SerAsn 384	
RESULT 14			
US-10-311-879-28			
; Sequence 28, Application US/10311879			
; Publication No. US20030186275A1			
; GENERAL INFORMATION:			
; APPLICANT: University of Sheffield			
; TITLE OF INVENTION: Antigenic Peptides			
; FILE REFERENCE: toxin			
; CURRENT APPLICATION NUMBER: US/10/311,879			
; CURRENT FILING DATE: 2003-03-17			
; NUMBER OF SEQ ID NOS: 32			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 28			
; LENGTH: 2659			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-10-311-879-28			
Alignment Scores:			
Pred. No.:	3,71e-05	Length:	2659
Score:	166.50	Matches:	100
Percent Similarity:	37.40%	Conservative:	90
Best Local Similarity:	19.69%	Mismatches:	209
Query Match:	7.16%	Indels:	109
DB:	12	Gaps:	18
US-09-147-052-1 (1-1371) x US-10-311-879-28 (1-2659)			
QY	49	TATGGTACGAACATCATCTCCGAGTACC	75
Db	817	HisSerValSerSerAsnProSerThrLeuThrAlaProAlaAlaHisThrValAsnThr	836
QY	76	-----CAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGCTC	111

Db 1174 LeuAspGluGluIleAsnLysSerValThrThraepGlyMetThrGlnSerSerIleGln 1193
QY 1039 TTCCTTTAAGTTCGGTTAAAGCAGCTGATGCTAATAACAGTTGGATTACAATAACATA 1098
Db 1194 AlaTyrGluAsnAlaLysArgAlaGlyGlnThrGluSerThrAsnAlaGlnAsnValIle 1213
QY 1099 AATAATGGAATGTTCAACAAGTTGAGTTGCCACTTCA----- 1137
Db 1214 AsnAsnGlyAspAlaThrAspGlnGlnIleAlaLysThrLysValGluGluLys 1233
QY 1138 ---ACTAGTCAAAATAACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTT 1194
Db 1234 TyrAsnSerLeuLysGlnAlaIleAlaGlyLeuThrProAspLeuAlaProLeuGlnThr 1253
QY 1195 GCTAAATCGTTTATCATAGTTTAAAGTTGGCCAAAACAAATGGAATTAAAGTTTCCA 1254
Db 1254 AlaLysThrGlnLeu-----GlnAsnAspIleAsp-----GlnPro 1265
QY 1255 ACGGTGAAGGAATATGATAAAGTTGCGCCAAATGATGGCAACATTATCTTAGCTCA 1314
Db 1256 ThrSerThrThrGlyMetThrSerAlaSerIleAlaAlaPheAsnGluLysLeuSerAla 1285
QY 1315 AATGAAATAATGCTGATAAGATC 1338
Db 1286 AlaArgThrLysIleGlnGluIle 1293

RESULT 15

US-09-815-242-5885
; Sequence 5885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5885
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: *Staphylococcus aureus*
US-09-815-242-5885

Alignment Scores:
Pred. No.: 6,08e-05 Length: 1029
Score: 163.00 Matches: 91
Percent Similarity: 38.61% Conservative: 53
Best Local Similarity: 24.40% Mismatches: 146
Query Match: 7.01% Indels: 83
DB: 9 Gaps: 19

US-09-147-052-1 (1-1371) x US-09-815-242-5885 (1-1029)
QY 235 ACCCAATTTAGAAGCAGCGCGAATGGAGTTAAACAGATCTAAATCAATGCTAAAGCGATGACA 294
Db 73 ThrAspLeuAsnThrAlaMetGlyAsnLeuGlnGlyAlaIleAsnAspGluGlnThrThr 92
QY 295 TTAGCTTCACACAGAGACTATGCCAAGATTGAAGCTAGTATTATCATCTGCTATATAGTAA 354
Db 93 LeuAsnSer---GlnAsnTyrGlnAspAlaThrProSerLysLysThrAlaTyrThrAsn 111
QY 355 GCT---GAAACAGTTAAACAATACCTTAATGCACATTTAGAACACACTAAAATGCTAAA 411
Db 112 AlaValGlnAlaLysAspIleLeuAsnLysSerAsnGlyGlnAsnLys---ThrLys 130
QY 412 ACTAATTTAGAAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAGTAA 471
Db 131 AspGlnValThrGluAlaMetAsnGlnValAsnSerAlaLysAsnAsnLeuAspGly--- 149
QY 472 CACCCAAATTTAGTTGAAGCATACAAAGCATATAAACCACTTTAGAACACAGCTGCTACT 531
Db 150 -----ThrArgLeuLeuAspGlnAlaLysGlnThrAlaLysGlnGlnLeuAsn 165
QY 532 AACCTTGAAGCTTTGTCATCACTGCTTATATCAATTCGCAATAATTTAGTGCATCTA 591
Db 166 AsnMetThrHisLeuThrThrAlaGlnLysThrAsnLeuThrAsnGln----- 181
QY 592 TACAATAAAGCTAGTAGTTTAATAACTAAACACTAGATCCCATCAATGAGGGAACGCTT 651
Db 182 -----IleAsnSerGlyThrThr 187
QY 652 TTAGATTCCTAATGAGATTACTACAGCTAATTAAGAAAT---ATTAATAATACGTTATCAACT 708
Db 188 ValAlaGlyValGlnThrValGlnSerAsnAlaAsnThrLeuAspGlnAlaMetAsnThr 207
QY 709 ATTAATGAACAAAGACATAAGCTGATGCA-----TTATCTAATAGTTTATT----- 756
Db 208 LeuArgGlnSerIleAlaAsnLysAspAlaThrLysAlaSerGluAspTyrValAspAla 227
QY 757 -----AAAAAGTGATTCAAAATAATCAACAAAGTTTGTAGGCACTTTTACAAAC 807
Db 228 AsnAsnAspLysGlnThrAlaTyrAsnAsnAlaValAlaAlaGluThrIleIleAsn 247
QY 808 GCTAATGTTCAACCTTCA-----AACTACAGTTTGTGCTTTTGTAGTGTGATACACCC 864
Db 248 AlaAsnSerAsnProGluMetAsnProSerThrIleThrGlnLysAlaGlu-----Gln 265
QY 865 GTCATTTATATATATGCAAGAGGACCGTTTGGATGCTGATGAA----- 909
Db 266 ValAsn-----SerSerLysThrAlaLeuAsnGlyAspGluAsnLeuThrAlaAla 282
QY 910 ---CCTTCAAGTAGAATTCCTTGCAAACACG---AATAGTATACAGAT----- 951
Db 283 LysGlnAsnAlaLysThrTyrLeuAsnThrLeuThrSerIleThrAspAlaGlnLysAsn 302
QY 952 -----GTTCTCTGGATTTATAGT-----TTAGCTGAACAAACACGAGTACCAA 996
Db 303 AsnLeuIleSerGlnIleThrSerAlaThrArgValSerGlyValAspThrValLysGln 322
QY 997 TTTAGTTTTCACACTATGGTCCCATCACTGGTATTATATATTTCCCTTATAGTTGGTT 1056
Db 323 AsnAla----- 324
QY 1057 AAAGCAGCTGATGCTAATAACGTTGGATTACAATAACAAATTAATAATGAAATGTTCAA 1116
Db 325 GlnHisLeuAspGlnAlaMetAlaSerLeuGlnAsnGlyIleAsnAsn----- 340
QY 1117 CAAGTTGAGTTGGCAGCTTCAACTAGTGCATAATATACTACAGCTAATCCAACTCCACGA 1176
Db 341 GluSerGlnValLysSerSerGluLysTyrArgAspAlaAspThrAsnLysGlnGlnGlu 360
QY 1177 GTTGATGAG---ATTAAAGTTGCTAAATCGTTTAA-----TCAGGTTTAAAGTTT 1224
Db 1177 GTTGATGAG---ATTAAAGTTGCTAAATCGTTTAA-----TCAGGTTTAAAGTTT 1224

Db 361 TyrAspAsnAlaIleThrAlaAlaLysAlaIleLeuAsnLysSerThrGlyProAsnThr 380
 QY 1225 GCCAAACACAAATCGAATTAGTGTTCACACGGGTG 1261
 Db 381 AlaGlnAsnAlaValGluAlaAlaLeu-GlnArgVal 392

Search completed: October 8, 2003, 19:26:24
 Job time : 109.102 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:41:47 ; Search time 21.5855 Seconds
(without alignments)
5374.731 Million cell updates/sec

Title: US-09-147-052-1

Perfect score: 2326

Sequence: 1 atgcaattatttagcgga.....gaccgcgtacattttataa 1371

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO/spool/US09147052/runat_08102003_154342_29880/app_query.fasta_1.5980
-DB=Issued Patents_AA -QEXT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052@cgn_1.1.126 @runat_08102003_154342_29880 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PTCUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	82.1	615	2	US-08-525-742-10
2	1856	79.8	610	2	US-08-525-742-8
3	1689	72.6	368	1	US-08-525-742-4
4	1663	71.5	368	1	US-08-185-851A-4
5	807	34.7	661	2	US-08-525-742-2
6	314	13.5	865	1	US-07-803-633A-13
7	185.5	8.0	10182	4	US-09-134-001C-3159
8	174.5	7.5	1041	1	US-08-220-151-4
9	174.5	7.5	1041	1	US-08-413-118-4
10	174.5	7.5	1041	3	US-08-473-446-4
11	173.5	7.5	933	3	US-08-293-728-2
12	173.5	7.5	933	3	US-09-421-868-2

13	154.5	6.6	682	4	US-08-836-687B-37	Sequence 37, Appl
14	146.5	6.3	1073	3	US-09-541-782-6	Sequence 6, Appl1
15	146.5	6.3	1073	4	US-09-723-820-6	Sequence 6, Appl1
16	144.5	6.2	930	4	US-09-134-001C-5314	Sequence 5314, Ap
17	140.5	6.0	1002	4	US-09-268-347-24	Sequence 24, Appl
18	140	6.0	2137	4	US-09-134-001C-4463	Sequence 4463, Ap
19	139	6.0	1004	4	US-09-268-347-30	Sequence 30, Appl
20	136	5.8	943	3	US-08-911-321-4	Sequence 4, Appl1
21	136	5.8	1098	3	US-08-923-992A-8	Sequence 8, Appl1
22	133.5	5.7	903	3	US-08-804-439A-22	Sequence 22, Appl
23	133.5	5.7	903	3	US-08-720-229-22	Sequence 22, Appl
24	133.5	5.7	904	3	US-08-632-537-1	Sequence 1, Appl1
25	133.5	5.7	904	5	PCI-US96-05316-1	Sequence 1, Appl1
26	133.5	5.7	904	6	5244792-4	Patent No. 5244792
27	133.5	5.7	1073	4	US-09-206-942-49	Sequence 49, Appl
28	133.5	5.7	1079	4	US-09-206-942-47	Sequence 47, Appl
29	132.5	5.7	903	1	US-08-220-151-8	Sequence 8, Appl1
30	132.5	5.7	903	1	US-08-413-118-8	Sequence 8, Appl1
31	132.5	5.7	903	3	US-08-473-446-8	Sequence 8, Appl1
32	131.5	5.7	1164	3	US-08-923-992A-2	Sequence 2, Appl1
33	131	5.6	1104	3	US-08-923-992A-4	Sequence 4, Appl1
34	129.5	5.6	1095	4	US-09-206-942-45	Sequence 45, Appl
35	129.5	5.6	1101	4	US-09-206-942-43	Sequence 43, Appl
36	129	5.5	1095	4	US-09-206-942-69	Sequence 69, Appl
37	129	5.5	1536	1	US-08-038-682-2	Sequence 2, Appl1
38	129	5.5	1536	1	US-08-302-832-2	Sequence 2, Appl1
39	129	5.5	1536	2	US-08-530-198-2	Sequence 2, Appl1
40	129	5.5	1536	2	US-08-469-880-2	Sequence 2, Appl1
41	129	5.5	1536	2	US-08-728-470-2	Sequence 2, Appl1
42	129	5.5	1536	2	US-08-617-697-2	Sequence 2, Appl1
43	129	5.5	1536	3	US-08-719-641-2	Sequence 2, Appl1
44	129	5.5	1536	4	US-09-206-942-67	Sequence 67, Appl
45	129	5.5	1833	4	US-08-621-944A-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hiroko

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigemi

; APPLICANT: Takahashi, Kiyohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; TITLE OF INVENTION: AS USE THEREOF

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; ADDRESS: NAUGHTON

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006

; COMPUTER READABLE FORM:

; COMPUTER TYPE: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 05-074139
 FILING DATE: 31-MAR-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: JP 05-245625
 FILING DATE: 30-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP94/00541
 FILING DATE: 31-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: McLealand, Le-Nhung
 REGISTRATION NUMBER: 31,541
 REFERENCE/DOCKET NUMBER: 950811
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-659-2930
 TELEFAX: 202-8870357
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 615 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-525-742-10

Alignment Scores:
 Pred. No.: 9,76e-151 Length: 615
 Score: 1910.00 Matches: 376
 Percent Similarity: 98.96% Conservative: 3
 Best Local Similarity: 98.17% Mismatches: 4
 Query Match: 82.12% Indels: 0
 DB: 2 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-525-742-10 (1-615)

QY 190 TGTATGCTATTACTAAAGAGTCAAAACCCAAATAATGCCAAACCCAAATTAAGACGA 249
 Db 27 CysMetSerIleThrLysLysAspAlaAsnProAsnGlyGlnThrGlnLeuGlnAla 46
 QY 250 GCGGAAATGGATTAACAGACTCAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAA 309
 Db 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
 QY 310 GACTATGCCAGATTAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTTAAAC 369
 Db 67 AspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGluThrValAsn 86
 QY 370 AATAACCTTAATGCAACATTAGAACACTAAATGCTAAACCTAAATTTAGTAATCAGCC 429
 Db 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 QY 430 ATCAACCAAGCTAATACGATAAAACGACTTTTGTATGATAATGAACACCCAAATTTAGTTGA 489
 Db 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
 QY 490 GCATACAAACACTAAACCACTTTAGAACACGCTGCTACTACCTTGAAGCTTTGTCTCA 549
 Db 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
 QY 550 TCAACTGCTTATATCAATTCGCAATTAATTTAGTGGATCTATACATAAAGCTAGTAGT 609
 Db 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAlaSerSer 166
 QY 610 TTAATACTAAACACTAGATCCACTAAATGGGGAAACGCTTTAGATTCTAATGAGATT 669
 Db 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
 QY 670 ACTACAGCTAATAAGAAATTAATAATAGCTTATCACTATTATTAATGAACAAAAGCTAAT 729
 Db 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
 QY 730 GCTGATGCTATTCTAATAGTTTATTAAAGAGTCAATCAAAATTAATCAACAAAGCTTTT 789
 Db 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226

QY 790 GTAGGACTTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGTAGT 849
 Db 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
 QY 850 GCTGATGTAAACACCCGCTCAATTAATAATATCAAGAGGACCGTTTGGAAATCGTGATCAA 909
 Db 247 AlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTrpAsnGlyAspGlu 266
 QY 910 CCTTCAAGTAGAATTTCTTGGCAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGT 969
 Db 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrpIleTyrSer 286
 QY 970 TTAGCTGGAACAAACACAGCAAGTACCAATTTAGTTTATAGCAACTATGGTCCATCACTGGT 1029
 Db 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
 QY 1030 TATTTATATTTCCCTTATAGTTTAAAGCAGCTGATGCTAATAACGTTGGATTACAA 1089
 Db 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGln 326
 QY 1090 TACAAATTAATTAATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAAT 1149
 Db 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
 QY 1150 AATACTACAGCTAATCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTGA 1209
 Db 347 AsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLysIleValLeu 366
 QY 1210 TCAGGTTTAAGATTTGCCAAACACAAATCAATTAAGTTGTTCCACCGGTGAAGGAAT 1269
 Db 367 SerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGlyGluGlyAsn 386
 QY 1270 ATGAATAAAGCTGCGCCAAATGATGCGCAACTTTTACGCTCAAAATGAAAATAATGCT 1329
 Db 387 MetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGluAsnAla 406
 QY 1330 GATAAGATC 1338
 Db 407 AspLysIle 409

RESULT 2

US-08-525-742-8
 ; Sequence 8, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Sakai, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hiroko
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemi
 ; APPLICANT: Takahashi, Kiyohito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSEE: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742

FILED DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-8

Alignment Scores:
Pred. No.: 2,94e-146 Length: 610
Score: 1856.00 Matches: 365
Percent Similarity: 98.17% Conservative: 10
Best Local Similarity: 95.55% Mismatches: 7
Query Match: 79.79% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-525-742-8 (1-610)

QY 190 TGTATGCTATTACTAAAGAGTCAACCAATATGTCACCAATAGCCAAATAGGACCA 249
DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGluAla 46
QY 250 GCGGAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCTTCACTACAA 309
DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGln 66
QY 310 GACTATGCCAAGATGAGCTAGTTATCATCTGCTATAGTGAAGCTGAACAGTAAAC 369
DB 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
QY 370 AATACCTTATGCAACATTAAGCACTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 429
DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGATTAACAGCTTTTGTATGATGAACACCAATTTAGTTGAA 489
DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
QY 490 GCATACAAAGCACTAAACCACTTTAGAACAGCTGCTACTACCTTTGAAGGTTGTGCA 549
DB 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSer 146
QY 550 TCAACTGCTTATATCAATTCGCAATATTTAGTGGATCTATACATAAAGCTAGTAGT 609
DB 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSer 166
QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTTAATGAGATT 669
DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIle 186
QY 670 ACTACAGCTAATAGATTAATTAATACGTTATCAACTATTAATGAACAAAGCACTAAT 729
DB 187 ThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GCTGATGCATTAATAGTTATTTTAAATAAAGCTGATTCATAATATGAACAAAGTTT 789

Db 207 AlaAspAlaLeuAlaAsnSerPheIleLysGluValIleGlnAsnAsnLysGlnSerPhe 226
QY 790 GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTACT 849
Db 227 ValGlyMetPheThrAsnThrAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
QY 850 GCTGATGTAAACCCGCTCAATTAATATATATCAAGAGGACCGTTTGGAAATGGTGAATA 909
Db 247 AlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTyrAsnGlyAspGlu 266
QY 910 CCTCAAGTAGAATTTGCAACACAGATAGTATACAGATGTTCTTGATTTATATAGT 969
Db 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerIleTyrSer 286
QY 970 TTAGCTGCACAAACACAGCAATGTTTCAACCTTATGTTTACCACTATGGTCCATCAACTGCT 1029
Db 287 LeuSerGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
QY 1030 TATTATATTTCCCTTATATAGTTTAAAGCAGCTGATGATTAACGTTGGATGATTAACA 1089
Db 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaSerAsnValGlyLeuGln 326
QY 1090 TACAAATTAATATGGAATGTTCAACAGCTTGGTTGGCTTCACTTCACTAGTCAAT 1149
Db 327 TyrLysLeuAsnAsnGlyAsnValGlnProValGluPheAlaThrSerThrSerAlaAsn 346
QY 1150 AATACCTACAGCTAATCCAACTCCAGCAGTTGATGATGATTAAGTTGCTAAATCGTTT 1209
Db 347 AsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLysIleValLeu 366
QY 1210 TCAGTTTAAAGATTGGCCAAACACAAATCGAATTAAGTTTCCACGGGTGAAGGAAT 1269
Db 367 SerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGlyGluArgAsn 386
QY 1270 ATGATTAAGTTGGCCCAATGATGGCACTATGCTTACCTTACCTCAATGAAATATGCT 1329
Db 387 MetAsnLysValAlaProMetIleGlyAsnMetTyrIleThrSerSerAsnAlaGluAla 406
QY 1330 GATAAG 1335
Db 407 AsnLys 408

RESULT 3

US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroo
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigeml
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELLAND &
; ADDRESS: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeiland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-4

Alignment Scores:
Pred. No.: 1,89e-132 Length: 368
Score: 1699.00 Matches: 335
Percent Similarity: 98.54% Conservative: 3
Best Local Similarity: 97.67% Mismatches: 4
Query Match: 72.61% Indels: 1
DB: 2 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-525-742-4 (1-368)

QY 190 TGTATGCTATTACTAAAGATGCAACCCCAATTAATGGCCAAACCAATTAGAGCA 249
Db |||||||
QY 27 CysMetSerIleThrLysLysAspAlaAsnProAsnGlyGlnThrGlnLeuGlnAla 46
Db |||||||
QY 250 GCGGGAATGAGTTACAGACTTAATCAATGCTTAAAGCGATGACATAGCTTCACTACAA 309
Db |||||||
QY 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
Db |||||||
QY 310 GACTATGCCAAGATTAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 369
Db |||||||
QY 67 AspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGluThrValasn 86
QY 370 AATAACCTTAATGCAACATTAGAACAACTAAATAATGGCTAAACCTTAATTTAGAAATCAGCC 429
Db |||||||
QY 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGATTAACGACTTTTGATAATGAACACCCCAATTTAGTTGAA 489
Db |||||||
QY 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
QY 490 GCATACAAAGCACTAAACCACTTTAGAACACCTGCTACTAACTTGAAGTTTGCA 549
Db |||||||
QY 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
QY 550 TCAACTGCTTATTAATCAATTCGCAATATTTAGTGGATCTATACAATAAAGCTAGTAGT 609
Db |||||||
QY 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
QY 610 TTAATACTAAACACTAGATCCACTAATAATGGGGAACGCTTTTATAGATTTAATGAGATT 669
Db |||||||
QY 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
QY 670 ACTACAGCTAATAACAATTAATAATACGTTATCACTATTAATGACAAAGCAATTAAT 729
Db |||||||

Db 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GCTGATGCATATCTAATAGTTTTTATTAAAGAGTGATTCAAAATAATGAACAAAGTTTT 789
Db |||||||
QY 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226
Db |||||||
QY 790 GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGT 849
Db |||||||
QY 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
Db |||||||
QY 850 GCTGATGTAACACCCGTCATTAATATGCAAGAGGACCGTTTGGAAATGCTGATGAA 909
Db |||||||
QY 247 AlaAspValThrProValAsnTyrLysTyrAlaArgThrValIlePasnGlyAspGlu 266
QY 910 CCTTCAGTAGAATTTCTTGCACACAGATAGTATCACAGATGTTTCTTGGATTTATAGT 969
Db |||||||
QY 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyrPheTyrSer 286
QY 970 TTAGCTGGAAACAAACACGAGTACCAATTTAGTTTGTAGCAACTATGTCCTCATCAACTGGT 1029
Db |||||||
QY 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
QY 1030 TATTTATATTCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGATGATACAA 1089
Db |||||||
QY 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGln 326
QY 1090 TACAAATTAATAATGGAATGTTCAACAGTTGAGTTTGGCTTCCCACTTCAACTAGTCAAT 1149
Db |||||||
QY 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
QY 1150 AATACACAGCTAATCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTAA 1209
Db |||||||
QY 347 AsnThrThrAlaAsnProThr-GlnGlnLeuMetArgLeuLysLeuLysSerPheTy 366
QY 1210 TCAGGTT 1216
Db 366 rGlnVal 368

RESULT 4
US-08-185-851A-4
Sequence 4, Application US/08185851A
Patent No. 5489430
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Fujisawa, Ayumi
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
TITLE OF INVENTION: Well As Vaccines Utilizing the Same
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Armstrong, Westerman, Hattori, McLeiland &
ADDRESSER: Naughton
STREET: 1725 K Street, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
SOFTWARE: ASCII from Word Perfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,851A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-659-2930

TELEFAX: 202-887-0357

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-185-851A-4

Alignment Scores:

Pred. No.: 2,72e-130 Length: 368
 Score: 1663.00 Matches: 333
 Percent Similarity: 97.96% Conservative: 3
 Best Local Similarity: 97.08% Mismatches: 6
 Query Match: 71.50% Indels: 1
 DB: 1 Gaps: 0

US-09-147-052-1 (1-1371) x US-08-185-851A-4 (1-368)

```

QY 190 TGTATGCTATTACTAAAGAGTGCACAAACCCCAAAATTAATGGCCAAACCCCAATTAGAGCA 249
DB 27 CysmetSerilehrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
QY 250 GCGGAATGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCCACTACAA 309
DB 47 AlaArgMetGluLeuThrAspLeuLeuAlaAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
QY 310 GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTAAC 369
DB 67 AspTrpAlaLysIleGluAlaSerLeuSerSerAlaTrpSerGluAlaGluThrValAsn 86
QY 370 AATAACCTTAAATGCAATAGAACACTAACTAAATGGCTAAATTAATTTAGATCAGCC 429
DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGGATAAAGACACTTTTGATTAATGAACACCCAAATTTAGTTGAA 489
DB 107 IleAsnGlnAlaAsnThrAspLysThrPheAspAsnGluHisProAsnLeuValGlu 126
QY 490 GCATACAAAGCACTAAACACCTTTAGAACACGCTACTACCTGAAGCTTGAAGTTCTCA 549
DB 127 AlaTrpLysAlaLeuLysThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
QY 550 TCAACTGCTTATATCAAAATCGCAATTAATTTAGTGGATCTATACAAATAAAGCTAGTAGT 609
DB 147 SerThrAlaTrpAsnGlnIleArgAsnAsnLeuValAspLeuTyAsnAlaSerSer 166
QY 610 TTATAACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCTTAATGAGATT 669
DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
QY 670 ACTACACTTAATAGAGATATTAATATACCTTATCAACTATTATGAACAAAGACCTAAT 729
DB 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GCTGATGCTATATAGTATTTATTAATAAAGTGATTCAAATAATGAACAAAGTTT 789
DB 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226
QY 790 GTAGGACTTTTCAAAACGGTAAATGTTCAACCTTCAACACTACAGTTTGTGCTTTTGTAGT 849
DB 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTrpSerPheValAlaPheSer 246
QY 850 CCTGATCAACACCCGCTCAATTAATATATATCCAGAAGGACCGTTTGGATGGTCAGCAA 909
DB 247 AlaAspValThrProValAsnTrpLysThrAlaArgArgThrVal***AsnGlyAspGlu 266
QY 910 CCTCAAGTGAATCTTGCACACCAACCAATAGTATCACAGATGTTCTTGGATTATAGT 969
DB 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSer***IleTrpSer 286

```

```

QY 970 TTAGCTGGAACAACACGAGTACCAATTTAGTTTACCAACTATGTCATCAACTGGT 1029
DB 287 LeuAlaGlyThrAsnThrLysTrpGlnPheSerPheSerAsnTrpGlyProSerThrGly 306
QY 1030 TATTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTATATAACGTTGGATTACAA 1089
DB 307 TyrLeuTrpPheProTrpLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGln 326
QY 1090 TACAAATTAATAATGAAATGTTCAACAAGTTGAGTTTGGCCACTTCAACTAGTGCAAAAT 1149
DB 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
QY 1150 AATACACAGCTAATCCAACTCCAGCAGTGTGATGAGATTAAGTTGCTAAAATCGTTTAA 1209
DB 347 AsnThrThrAlaAsnProThrGlnGlnLeuMetArgLeuLysLeuLysSerPheTy 366
QY 1210 TCAGGTT 1216
DB 366 rGlnVal 368

```

RESULT 5

US-08-525-742-2

Sequence 2, Application US/08525742

Patent No. 5871742

GENERAL INFORMATION:

APPLICANT: Saito, Shuji

APPLICANT: Ohkawa, Setsuko

APPLICANT: Saeki, Sakiko

APPLICANT: Ohsawa, Ikuroh

APPLICANT: Funato, Hiroko

APPLICANT: Iritani, Yoshikazu

APPLICANT: Aoyama, Shigemasa

APPLICANT: Takahashi, Kiyoochito

TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

TITLE OF INVENTION: AS USE THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

STREET: 1725 K Street, Suite 1000

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/525,742

FILING DATE: 25-SEP-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 05-074139

FILING DATE: 31-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 05-245625

FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP94/00541

FILING DATE: 31-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: McLealand, Le-Nhung

REGISTRATION NUMBER: 31,541

REFERENCE/DOCKET NUMBER: 950811

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-659-2930

TELEFAX: 202-8870357

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

```

US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; City: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:

```


SEQUENCE CHARACTERISTICS:
 LENGTH: 865 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-803-633A-13

Alignment Scores:

Pred. No.: 2,74e-18 Length: 865
 Score: 314.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.50% Indels: 0
 DB: 1 Gaps: 0

US-09-147-052-1 (1-1371) x US-07-803-633A-13 (1-865)

QY 1 ATGCACATTTTAGCGGAATGTCATATTTTCTTATAGTTATCTATATGTTACGAAAC 60
 DB 1 MethH1sTyPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyGlyThrAsn 20
 QY 61 TCATCTCGAGTACCCAAAATGTGACATCAAGAGAGTGTTCGAGCGTCCAGTTGTCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerValGlnLeuSer 40
 QY 121 GAGGAGAGTCTACGTTTATCTTTGTCCTCCACCAGTGGTTCACCCGTGATCGTGCTA 180
 DB 41 GluGluGluSerThrPheTyLeuCysProProValGlySerThrValIleArgLeu 60
 QY 181 GAA 183
 DB 61 Glu 61

RESULT 7

US-09-134-001C-3159
 Sequence 3159, Application US/09134001C
 Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 3159
 LENGTH: 10182
 TYPE: PNT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3159

Alignment Scores:

Pred. No.: 2,12e-07 Length: 10182
 Score: 185.50 Matches: 122
 Percent Similarity: 32.37% Conservative: 69
 Best Local Similarity: 20.68% Mismatches: 202
 Query Match: 7.98% Indels: 197
 DB: 4 Gaps: 23

US-09-147-052-1 (1-1371) x US-09-134-001C-3159 (1-10182)

QY 58 AACTCATCTCGAGTACCCAAAATGTGACA---TCAAGAGAGTGTGTTTCGAGCGTCCAG 114
 DB 3981 AsnSerAsnProSerValAsnGluValAlaGlnAlaLeuGlnLysValGluAlaValGln 4000
 QY 115 TTGTCGTGAGGAGTCTACGTTTATCTTTTGTCTCCGCCACCGTGGTTCAACCGTGATC 174
 DB 4001 LeuLysValAsnAspAlaIleHisIleLeu----- 4010

QY 175 CGTCTAGAAATTCGGCTGTATGTCTATTACTAAAAAAGATGCAAAACCCAAATAATGCGCAA 234
 DB 4011 -----GlnAsnLysGluAsn---Asn 4016
 QY 235 ACCCAATTAGAAGCAGCGCATGAGTTAAACAGATCTAACTCAATGCTTAAGCGGATGACA 294
 DB 4017 SerAlaLeuValThrAlaLysAsnGlnLeuGlnSerValAsnAspGlnProLeuThr 4036
 QY 295 TTAGCTTCACACAAAGC-----TATGCCAAGATTGAAGCTAGTTTATCATCT 342
 DB 4037 ThrGlyMetThrGlnAspSerIleAsnAsnTyrgluAlaLysArgAsnGluAlaGlnSer 4056
 QY 343 GCTTATAGTGAAGCTGAACACAGTTAAACAATAACCTTAATGACACATTGAACAACTAAA 402
 DB 4057 AlaIleArgAsnAlaGluAlaValIleAsnAsnGlyAspAlaThrAlaLysGlnIleSer 4076
 QY 403 ATGGCTAAA----- 411
 DB 4077 AspGluLysSerLysValGluGlnAlaLeuAlaHisLeuAsnAspAlaLysGlnGlnLeu 4096
 QY 412 -----ACTAATTTAGAAATCAGCCATCAACCAAGCTAAATACGGATAAAACGACT 459
 DB 4097 ThrAlaAspThrThrGluLeuGlnThrAlaValGlnGlnLeuAsnArgArgGlyAspThr 4116
 QY 460 TTTGATATGACACCCCAATTTAGTTGACGATACAAAGCACTAAAACCACTTTAGAA 519
 DB 4117 ---AsnAsnLysLysProArgSerIleAsnAlaTyrglnLysAlaIleGlnSerLeuGlu 4135
 QY 520 CAACGTGCTACTAACTT---GAAGTTTGTCTCACTCACTGCTTATATCAATCAATTCGC--- 573
 DB 4136 ThrGlnIleThrSerAlaLysAspAsnAlaAsnAlaValIleGlnLysProIleArgThr 4155
 QY 574 -----AATAATTTAGTGCATCTATACAAATAAAGCTAGTTTATATACTAAA 621
 DB 4156 ValGlnGluValAsnAsnAlaLeuGlnGlnValAsnGlnLeuAsnGlnGlnLeuThrGlu 4175
 QY 622 ACA-----CTAGATCCACTAAATGGGGGAAACGCTTTTAGATCTTAATGAGATTACT 672
 DB 4176 AlaIleAsnGlnLeuGlnProLeuSerAsnAsnAspAlaLeuLysAlaAlaArgLeuAsn 4195
 QY 673 ACAGCTAATAAGATTAATAATACGTTATCACTTATCACTTATTAATGAACAAAGACTAATGCT 732
 DB 4196 LeuGluAsnLys---IleAsnGlnThrValGlnThrAspGlyMetThrGlnGlnSerIle 4214
 QY 733 GATGCATTATCTAATAGTTTATTAATAAAGTGATTCAAAATAATGAACAAAGTTTGTGA 792
 DB 4215 GluAlaTyrglnAsnAla-----LysArgValAlaGlnAsnGluSerAsnThrAlaLeu 4232
 QY 793 GGGACTTTTACAAAC----- 807
 DB 4233 AlaLeuIleAsnAsnGlyAspAlaAspGluGlnGlnIleThrThrGluThrAspArgVal 4252
 QY 808 -----GCTAATGTTCAACCT 822
 DB 4253 AsnGlnGlnThrThrAsnLeuThrGlnAlaIleAsnGlyLeuThrValAsnLysGluPro 4272
 QY 823 TCAAACTACAGTTTGTGCTTTTAGTGTGCTGATGATACACCCGCTC----- 867
 DB 4273 LeuGluThrAlaLysThrAlaLeuGlnAsnAlaIleAspGlnValProSerThrAspGly 4292
 QY 868 -----AATTATAAATGTCAGAGGACCGCTTTGGTATCGTAT 906
 DB 4293 MetThrGlnGlnSerValAlaAlaAsnGlnLysLeuGlnIleAlaLysAsnGlnIle 4312
 QY 907 GAACCTTCAAGTAGAATTTCTTCAAC----- 933
 DB 4313 AsnThrIleAsnAsnValLeuAlaAsnAsnProAspValAsnAlaIleIleThrAsnLys 4332
 QY 934 -----ACGAATAGTATCA-----GAT 951
 DB 4333 AlaGluAlaGluArgIleSerAsnAspLeuThrGlnAlaLysAsnAsnLeuGlnValAsp 4352
 QY 952 GTTCTTGGATTATAGTTTA-----GCTGGAAACAAC 984

```

Db 4353 ThrGlnProLeuGluLysIleLysArgGlnLeuGlnAspGluIleAspGlnGlyThrAsn 4372
QY 985 ACG-----AAGTACCAATTTAGTTTATGCAACTATGCTCCATCA---ACTGGTTATTATA 1035
Db 4373 ThrAspGlyMetThrGlnAspSerValAspAsnTyrAsnSerLeuSerAlaAlaIle 4392
QY 1036 TATTTCCCTTATAGCTGTTAAAGCAGCTGATGCTATATACGTTGGATTACATACAAA 1095
Db 4393 IleGluLysGlyLysValAsnLysLeuLeuLysArgAsnProThrValGluGlnValLys 4412
QY 1096 TTAATAATGGAATGTTCAACAAGTT----- 1122
Db 4413 GluSerValAlaAsnAlaGlnGlnValIleGlnAspLeuGlnAsnAlaArgThrSerLeu 4432
QY 1122 ----- 1122
Db 4433 ValProAspLysThrGlnLeuGlnGluAlaLysAsnArgLeuGluAsnSerIleAsnGln 4452
QY 1123 -----GAGTTTGCCCACTCAACTAGTCAATATATATACTACAGCTATCCAACT 1170
Db 4453 GlnThrAspThrAspGlyMetThrGlnAspSerLeuAsnAsnTyrAsnAspLysLeuAla 4472
QY 1171 CCAGCAGTTGATGAGATT---AAAGTTGCTAAATCGTTTATCAGGTTTAAGATTGGC 1227
Db 4473 LysAlaArgGlnAsnLeuGluLysIleSerLysValLeu-----GlyGly 4487
QY 1228 CAAAACACAATC---GAATTAAAGTGTCCCAACGGTCAAGGAATATGAATAAGTTGCG 1284
Db 4488 GlnProThrValAlaGluIleArgGlnAsnThrAspGluAlaAlaHisLysGlnAla 4507
QY 1285 -----CCAATGATTGCCAATTTAT 1305
Db 4508 LeuAspThrAlaArgSerGlnLeuThrLeuAsnArgGluProTyrIleAsnHisIleAsn 4527
QY 1306 CTTAGCTCAATGAAATATGCTGATAAG 1335
Db 4528 AsnGluSerHisLeuAsnAlaGlnLys 4537

RESULT 8
US-08-220-151-4
; Sequence 4, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS

```

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-4

```

```

Alignment Scores:
Pred. No.: 1,07e-06 Length: 1041
Score: 174.50 Matches: 124
Percent Similarity: 31.66% Conservative: 71
Best Local Similarity: 20.13% Mismatches: 199
Query Match: 7.50% Indels: 222
DB: 1 Gaps: 24

```

```
US-09-147-052-1 (1-1371) x US-08-220-151-4 (1-1041)
```

```

QY 7 TATTTTAGCGCGAATTCATATTT----- 30
Db 22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCTATAT----- 51
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 52 -----GGTACGAACCTCATCTCCGAGTACCCAA 78
Db 82 ProArgArgThrValAlaThrProGluValGlyGlyThrProLysProThrThrAsp 101
QY 79 AATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGCTGTGAGGAAGAG----- 129
Db 102 ProThrAspMetSerMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
QY 130 ---TCTACGTTTATCTTTGTCGCCACCAGGTTCAACCGTATCGTCTAGAA--- 183
Db 122 ProSerThrPheTyrMetCysProProSerGlySerThrValArgLeuGluPro 141
QY 184 -----TTCGGCTGTATGCTATT 201
Db 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
QY 202 ACTAAAAGATGCCAACCCA----- 222
Db 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
QY 223 -----AATAATGGCCAAACCCAA 240
Db 182 MetThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
QY 241 TTAGAAGCAGCGCGAATCGAGTTACAGATCTAATCAATGCTAAAGCATGACATTAGCT 300
Db 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeu--- 220
QY 301 TCACTACAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTAGAGCTGAA 360
Db 221 -----SerLysAlaAsp 224
QY 361 ACAGTTACAATAACCTTTAATGCAACATTAAGAACAACTATAAAATGGCTATAAATTTA 420
Db 225 TyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArgGluLeu 244
QY 421 GAATCAGCCATCAACCAAGCTAATACGGATAAACAGACTTTTGTGATAATGAACACCCAAAT 480
Db 245 ProLeuLysProSerLysPheAsnThrProGlnSerArg-----GlyTyrPheHis----- 260
QY 481 TTAGTTGAAGCATACAAAGCAGCTAAACCACTTTAGNACACACGCTGCTACTAACCTTGA 540

```

```
Db 261 -----ThrTyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThr 277
QY 541 GCTTTGTCATCACTGCTTATAATCAATTCCTCAATAATTTAGTGGATCTATACATAAA 600
Db 278 AlaTpaAlaGlySerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIle 297
QY 601 GCTAGTATTTAATAACATAACACTAGATCCCAATAATGGGGACGCTTTAGATCT 660
Db 298 ProValSerGluIleThrAspThrIleAspLysPheGly-----LysCys 312
QY 661 AATGAGTACTACAGCTAATAAGAAATATTAATACGTTATCACTATTATTAACAA 720
Db 313 SerSerLysAlaThrTyrValArgAsn---AsnHisLysValGluAlaPheAsnGluAsp 331
QY 721 AAGACTAATGCTGAT-----GCATTATCTAATAGTATTTTATTAANAAGTG 765
Db 332 LysAsnProGlnAspMetProLeuIleAlaSerLysTyrAsnSerValGlySerLysAla 351
QY 766 ATTCAAAATATGAACAAGTTTCTGA-----GGG 795
Db 352 TrpHisThrThrAsnGluThrTyrThrLysIleGlyAlaAlaGlyPheHisSerGly 371
QY 796 ACTTTTACAAACGCTAATGTTCAACCT-----TCAAACTACAGTTTGTGCT 843
Db 372 ThrSerValAsnCysIleValGluGluValAlaArgSerValTyrProTyrAspSer 391
QY 844 TTTAGTGTGATGTAACACCCGCTCAATTAT-----TCAAACTACAGTTTGTGCT 873
Db 392 PheAlaIleSerThrGlyAspValIleHisMetSerProPhePheGlyLeuArgAspGly 411
QY 874 AAATATGCAGACGACCGCTGGATGGTGATGA-----909
Db 412 AlaHisValGluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyr 431
QY 910 -----CCTTCAAGTAGAATCTTGCACAA 933
Db 432 ProIleAspLeuAspThrArgLeuGlnLeuGlyAlaProValSerArgAsnPheLeuGlu 451
QY 934 AGCAATAGTATCAGATGTTCTTGG-----ATTATAGT 969
Db 452 ThrProHisValThr---ValAlaTrpAsnTrpThrProLysCysGlyArgValCysThr 470
QY 970 TTAGCTGGA-----ACAACACGAGTACCAATTTAGTTTACG 1008
Db 471 LeuAlaLysTrpArgGluIleAspGluMetLeuArgAspGluTyrGlnGlySerTyrArg 490
QY 1009 AACTATGGTCCATCACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAACGACGTGAT 1068
Db 491 -----PheThrValLysThrIleSerAlaThrPhe 500
QY 1069 GCTAATAACGTTGGATTACAAATTAATAATGGAATGTTTCAACAAAGTTGAGTTT 1128
Db 501 IleSerAsnThrSer---GlnPheGluIleAsn-----ArgIleArgLeuGlyAspCys 517
QY 1129 GCCACTTCACTAGTGCATAATTAATCTACAGCTATCCAACTCCAGCTGTTGATGATT 1188
Db 518 AlaThrLysGluAlaAlaGlu-----AlaIleAspArgIle 529
QY 1189 AAGTTGCTAAATCGTTTATCATGTTTAAAGTTTGGCAAAACACAACTCAATTAAGT 1248
Db 530 TyrLysSerLysTyrSerLysThrHisIleGlnThrGly-----ThrLeuGluThrTyr 547
QY 1249 GTTCAACGGGTGAAGGAATATGAATAAGTTGCGCCCAATGATTGGCAACATTTATCTT 1308
Db 548 LeuAlaArgGly---GlyPheLeuIleAlaPheArgProMetIleSerAsnGluLeuAla 566
QY 1309 AGCTCAATGAATAATAGCTGATAGATCCCGGGTACCGTCGAC 1354
Db 567 LysLeuTyrIleAsnGlu---LeuAlaArgSerAsnArgThrValAsp 581
RESULT 9
US-08-413-118-4
```

```
; Sequence 4, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-4
Alignment Scores:
Pred. No.: 1,07e-06 Length: 1041
Score: 174.50 Matches: 124
Percent Similarity: 31.66% Conservatives: 71
Best Local Similarity: 20.13% Mismatches: 199
Query Match: 7.50% Indels: 222
DB: 1 Gaps: 24
US-09-147-052-1 (1-1371) x US-08-413-118-4 (1-1041)
QY 7 TATTTTAGCGGAATTCATATTT-----30
Db 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCTATAT-----51
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 52 -----GGTACGAACCTCATCTCCGAGTACCCAA 78
Db 82 ProArgArgThrValAlaThrProGluValGlyThrProProLysProThrThrAsp 101
QY 79 AATGTGACATCAAGACAGTCTTTCGACGCTCCAGTGTCTCTGAGGAAGAG-----129
Db 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
```

```

QY 934 ACGAATAGTATCACAGATGTTCTTCGG-----ATTATAGT 969
Db      ||| :||| |||:|||
QY 452 ThrProHisValThr--ValAlaTrpAsnTrpThrProLysCysGlyArgValCysThr 470
Db      ||| :||| |||:|||
QY 970 TTAGCTGGA-----ACAAACAGGAAGTACCATAATTTAGTTTAGC 1008
Db      ||| ||| :||| |||:|||
QY 471 LeuAlaLysTrpArgGluLeuAspGluMetLeuArgaspGluTyrglnGlySerTyArg 490
Db      ||| ||| :||| |||:|||
QY 1009 AACATATGCTCCATCAACTGGTTATTATATATTCCTTATAAGTTGGTTAAAGCAGCTGAT 1068
Db      ||| ||| :||| |||:|||
QY 491 -----PheThrValLysThrIleSerAlaThrPhe 500
QY 1069 GCTAATAACGTGGATTACAATAACAANTTAATGGAATGTTCAACAAGTTGAGTTT 1128
Db      ||| ||| :||| |||:|||
QY 501 IleSerAsnThrSer---GlnPheGluIleasn-----ArgIleArgLeuGlyAspCys 517
Db      ||| ||| :||| |||:|||
QY 1129 GCCACTTCAAAGTCTAGTGCAAAATAATACTACAGCTAATCCAACCTCCAGCAGTTGATGAGATT 1188
Db      ||| ||| :||| |||:|||
QY 518 AlaThrLysGluAlaAlaGlu-----AlaIleAspArgIle 539
QY 1189 AAAGTTCTTAAATCGTTTTATCATCGTTTAAAGATTGGCCAAAACACAATCGAATTAAGT 1248
Db      ||| ||| :||| |||:|||
QY 530 TyrLysSerLysTyrSerLysThrHisIleGlnThrGly-----ThrLeuGluThrTyr 547
QY 1249 GTTCCAACGGGTGAAGAAATATGAATAAAGTTGGCGCAATGATGGCAACATTTATCTT 1308
Db      ||| ||| :||| |||:|||
QY 548 LeuAlaArgGly--GlyPheLeuIleAlaPheArgPrometIleSerSngluLeuAla 566
QY 1309 AGCTCAATGAANAATATGCTGTATAGATCCC CGGTACGTCGCAC 1354
Db      ||| ||| :||| |||:|||
QY 567 LysLeuTyrIleAsnGlu-LeuAlaArgSerAsnArgThrValAsp 581

RESULT 10
US-08-473-446-4
; Sequence 4, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/473,446
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO.: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-473-446-4

Alignment Scores:

Pred. No.: 1,07e-06 Length: 1041
Score: 174.50 Matches: 124
Percent Similarity: 31.66% Conservative: 71
Best Local Similarity: 20.13% Mismatches: 199
Query Match: 7.50% Indels: 222
DB: 3 Gaps: 24

US-09-147-052-1 (1-1371) x US-08-473-446-4 (1-1041)

```

QY 7 TATTTAGCGGAATTCGATATT----- 30
D 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
D 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATCTATAT----- 51
D 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 52 -----GGTAGGAATCATCTCCGAGTACCCAA 78
D 82 ProArgArgThrValAlaThrProGluValGlyGlyThrProProLysProThrThrAsp 101
QY 79 AATGTGACATCAAGAGAGTCTTTCGAGCGTCCAGTTGTCTGAGAGAG----- 129
D 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
QY 130 ---TCTACGTTTACTTGTCCCCACAGGGTTCACCGTGCATCGCTAGAA--- 183
D 122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
QY 184 -----TTCGGCTGTATGCTATT 201
D 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIle 161
QY 202 ACTAAAGAGATGCAACCCA----- 222
D 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIle 181
QY 223 -----AATAATGGCCAAACCCAA 240
D 182 MetThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
QY 241 TTAGAAGCAGCCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCT 300
D 202 ValProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeu--- 220
QY 301 TCACTACAGACTATGCCNAGATTGAGCTAGTTATCATCTGCTTATAGTGAAGCTGAA 360
D 221 -----SerLysAlaAsp 224
QY 361 ACAGTTAACAATAACCTTAATCAACATAGAACAACTAAAAATGGCTAAACTAATTTA 420
D 225 TyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArgGluLeu 244
QY 421 GAATCAGCCATCAACCAAGCTAATACGGATAAACGACTTTTGTATGAATGAACACCCAAAT 480
D 245 ProLeuLysProSerLysPheAsnThrProGlnSerArg-----GlyTrpHis 260
QY 481 TTAGTTGAGCATACAAAGCACTAAACACCACTTTAGAACACAGCTGCTACTAACCTTGA 540
D 261 -----ThrTyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThr 277
QY 541 GGTGTGTCATCACTGCTTATATCAATTCGCAATATTTAGTGATCTATACATATA 600

```

RESULT 11

US-08-293-728-2

; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: Mcdevitt, Damien L.

```

D 278 AlaTrpAlaGlySerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIle 297
QY 601 GCTAGTAGTTTAAATAACTAAACACACATAGATCCATTAATGGGGAAACGCTTTTAGATTCT 660
D 298 ProValSerGluIleThrAspThrIleAspLysPheGly-----LysCys 312
QY 661 AATGAGATTACTACAGCTAATAAGTAATATTAATACGTTATCACTAATTAATGAACAA 720
D 313 SerSerLysAlaThrTyrValArgAsn---AsnHisLysValGluAlaPheAsnGluAsp 331
QY 721 AAGACTAATGCTGAT-----GCATTATCTAATAGTTTATTAATAAAGG 765
D 332 LysAsnProGlnAspMetProLeuIleAlaSerLysTyrAsnSerValGlySerLysAla 351
QY 766 ATCAAAATAATGAACAAAGTTTGT-----GGG 795
D 352 TrpHisThrThrAsnGluThrTyrThrLysIleGlyAlaAlaGlyPheHisHisSerGly 371
QY 796 ACTTTTACAAACGCTAATGTTCAACCT-----TCAACTACAGTTTGTGCT 843
D 372 ThrSerValAsnCysIleValGluValAlaAspAlaArgSerValTyrProTyrAspSer 391
QY 844 TTTAGTGTGATGTAACACCCGCTCAATTAT--- 873
D 392 PheAlaIleSerThrGlyAspValIleHisMetSerProPhePheGlyLeuArgAspGly 411
QY 874 AATATGCAAGAGACCGTTCGTAATGCTGATGA----- 909
D 412 AlaHisValGluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrTyr 431
QY 910 -----CCTTCAAGTAGAATCTTCTGCAAC 933
D 432 ProIleAspLeuAspThrArgLeuGlnLeuGlyAlaProValSerArgAsnPheLeuGlu 451
QY 934 AGCAATAGTATCACAGATGTTCTTGG-----ATTATATCT 969
D 452 ThrProHisValThr---ValAlaTrpAsnTrpThrProLysCysGlyArgValCysThr 470
QY 970 TTAGCTGGA-----ACAAACACGAATACCAATTTAGTTTAGC 1008
D 471 LeuAlaLysTrpArgGluIleAspGluMetLeuArgAspGluTyrGlnGlySerTyrArg 490
QY 1009 AACTATGTCCTCACTGCTGTTATTTATTTCCCTTATATAGTTGTTAAACGACGCTGAT 1068
D 491 -----PheThrValLysThrIleSerAlaThrPhe 500
QY 1069 GCTAATAAGCTGGATTACAATAAATAATGGAATGTTCAACAAGTTGAGTTT 1128
D 501 IleSerAsnThrSer---GlnPheGluIleAsn-----ArgIleArgLeuGlyAspCys 517
QY 1129 GCACCTTCAACTAGTGCATAATAACTACAGCTAATCAACTCCACAGCTGATGAGATT 1188
D 518 AlaThrLysGluAlaAlaGlu-----AlaIleAspArgIle 529
QY 1189 AAGTTGCTAAATCGTTTATCAGGTTTAAGATTGGCCAAACACATCGAATTAAGT 1248
D 530 TyrLysSerLysTyrSerLysThrHisIleGlnThrGly-----ThrLeuGluThrTyr 547
QY 1249 GTTCCAAACGGTGAAGGAAATATGATAAAGTTGCGCAATGATTTGGCAACATTTATCTT 1308
D 548 LeuAlaArgGly---GlyPheLeuIleAlaPheArgPrometIleSerAsnGluLeuAla 566
QY 1309 AGCTCAAAATGAATAATGCTGATAAGATCCCGGGTACCGTCGAC 1354
D 567 LysLeuTyrIleAsnGlu-LeuAlaArgSerAsnArgThrValAsp 581

```

: TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene									
: FILE REFERENCE: 05344.105011									
: CURRENT APPLICATION NUMBER: US/08/293,728D									
: CURRENT FILING DATE: 1994-08-22									
: NUMBER OF SEQ ID NOS: 20									
: SOFTWARE: PatentIn Ver. 2.0									
: SEQ ID NO 2									
: LENGTH: 933									
: TYPE: PRM									
: ORGANISM: Staphylococcus aureus									
US-08-293-728-2									
Alignment Scores:									
Pred. No.: 1.26e-06 Length: 933									
Score: 173.50 Matches: 111									
Percent Similarity: 35.90% Conservative: 66									
Best Local Similarity: 22.52% Mismatches: 185									
Query Match: 7.46% Indels: 131									
DB: 3 Gaps: 23									
US-09-147-052-1 (1-1371) x US-08-293-728-2 (1-933)									
QY	4	CAC	TATTTAGGCGGAATGC	ATATTTTCCTTATAGTTATCTA	-----	48			
Db	9	His	AlaIleAlrGlyLysSer	IleGlyValAlaSerValLeuValGlyThrLeu	IIIIIIIIII	IleGly 28			
QY	49	TAT	GTCACGAAC	TCACTCT	-----	CCGAGTACCCAAATGACATCAAGAA 96			
Db	29	Phe	GlyLeuLeuSerSerLys	GluaAlaSerGluAsnSerValThrGlnSer	IIIIIIIIII	Asp 48			
QY	97	GTT	GTTTCGACGTC	CAGTCTCTGAGAGAGCT	TACGTTTTTATCTTTGTCCTCCCAACCA 156				
Db	49	Ser	AlaSerAsnGluSerLys	SerAsnAspSerSerValSerAlaAlaProLys	IIIIIIIIII	Thr 68			
QY	157	GTT	GCGTTCACCGCTGAT	CCGCTAGACTTCGGCTGTATGCTATTACT	ATAAAGATGCA 216				
Db	69	Asp	AspThrAsnVal	-----	-----	SerAspThrLysThrSerSer 80			
QY	217	AAC	CAATAATATG	CCCAACCAATTA	-----	GAAGCAGCGCGAATGGAGTTAACAGAT 270			
Db	81	Asn	ThrAsnAsnGlyGlu	ThrSerValAlaGlnAsnProAlaGlnGlnGlu	IIIIIIIIII	ThrGln 100			
QY	271	CTA	ATCAATGCTAA	AGCGATGACATAGCTT	CACTCAAGACTATGCCAAGATTGAAGCT 330				
Db	101	Ser	SerSerThrAsnAla	ThrThr	-----	-----	108		
QY	331	AGT	TATCATCTGCT	TATAGTGAAGCTGAACATTA	CAACCTTAACTTAATGCAACATTA 390				
Db	109	Glu	GluThrProValThr	GlyGluAlaThrThrThrThrThrAsnGlnAla	AsnThrPro 128				
QY	391	GAA	CACTAAATATG	CGTAAACTAAATTTAGATACAGCCATCAACCAAGCTAATACGAT 450					
Db	129	Ala	ThrThrGlnSerSer	AsnThrAsnAlaGluGluLeuValAsnGln	-----	-----	ThrSerAsn 147		
QY	451	AAA	ACGACTTTT	GATATGACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 510					
Db	148	Glu	ThrThrPhe	AsnAsp	-----	-----	thrAsnThrValSerSerValAsnSer	162	
QY	511	ACT	TTAGAACAACG	TGCTCATACCTTGAAGCTTTGTCATCACT	-----	-----	555		
Db	163	-----	ProGlnAsnSerThr	AsnAlaGluAsnValSerThrThrGlnAspThr	serThr 180				
QY	556	-----	GCTTTAATCAAA	TTTCGCAATAATTTAGTGGATCTATACAATAAA	-----	-----	600		
Db	181	Glu	AlaThrProSer	AsnAsnGluSerAlaProGlnSerThrAspAlaSer	AsnLysAsp 200				
QY	601	-----	-----	-----	-----	-----	GCTAGTAGTTTAATACT 618		
Db	201	Val	ValAsnGlnAlaVal	AsnThrSerAlaProArgMetArgAlaPhe	SerLeuAlaAla 220				
QY	619	AAA	CACTAGATCC	CTAAATGGGGAACGCTTTTAGATTCTAATGACATTACTACAGCT 678					
Db	221	Val	AlaAlaAspAla	ProAlaGlyThrAspIle	-----	-----	ThrAsnGlnLeuThrAsnVal 239		

Percent Similarity: 35.90% Conservative: 66
Best Local Similarity: 22.52% Mismatches: 185
Query Match: 7.46% Indels: 131
DB: 3 Gaps: 23

US-09-147-052-1 (1-1371) x US-09-421-868-2 (1-933)

```

Qy 4 CACTATTTTAGCGGAANTGCATATTTTCCCTTAGTATTCTA----- 48
Db 9 HIsAlaIleArGLysLysSerIleGlyValAlaSerValLeuValGlyThrLeuIleGly 28
Qy 49 TATGCTAGAACTCATC-----CGGAGTACCCAAATGTGACATCAAGAGAA 96
Db 29 PheGlyLeuLeuSerSerLysGluAlaAspAlaSerGluAsnSerValThrGlnSerAsp 48
Qy 97 GTTGTTTTCGACGCTCCAGTGTCTGAGGAAGAGTCTAGCTTTTATCTTTGTCCCCACCA 156
Db 49 SerAlaSerAsnGluSerLysSerAsnAspSerSerValSerAlaAlaProLysThr 68
Qy 157 GTGGTTCAACCGTGATCCGCTCTAGAAATTCGGCTGTATGTCTATTACTATAAAGATGCA 216
Db 69 AspAspThrAsnVal-----SerAspThrLysThrSerSer 80
Qy 217 AACCCAAATAATGGCAAAACCAATTA-----GAAGCAGCGCGAATGAGGTTAACAGAT 270
Db 81 AsnThrAsnAsnGlyGluThrSerValAlaGlnAsnProAlaGlnGluThrThrGln 100
Qy 271 CTAATCAATGCTAAAGCGATGACATAGCTTCATCACAAGACTATGCCAAGATTGAAGCT 330
Db 101 SerSerSerThrAsnAlaThr----- 108
Qy 331 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTTAAACAATTAACCTTAAATGCAACATTA 390
Db 109 GluGluThrProValThrGlyGluAlaThrThrThrThrAsnGlnAlaAsnThrPro 128
Qy 391 GAACAACCTAAAAATGGCTAAAACTAATTTAGAAATCAGCATCAACCAAGCTAATACGGAT 450
Db 129 AlaThrThrGlnSerSerAsnThrAsnAlaGluGluLeuValAsnGln---ThrSerAsn 147
Qy 451 AAAACGACTTTTGATTAATGACACCCAAATTTAGTTTGAAGCATACAAACGACTAAAAACC 510
Db 148 GluThrThrPheAsnAsp-----ThrAsnThrValSerSerValAsnSer----- 162
Qy 511 ACTTTAGAACAGCTGCTACTAAACCTGGAAGGTTTGTCTATCAACT----- 555
Db 163 -----ProGlnAsnSerThrAsnAlaGluAsnValSerThrThrGlnAspThrSerThr 180
Qy 556 -----GCTTTAATAACAAATTCGCAATAATTTAGTGGATCTATACATAATAA--- 600
Db 181 GluAlaThrProSerAsnAsnGluSerAlaProGlnSerThrAspAlaSerAsnLysAsp 200
Qy 601 -----GCTAGTAGTTTAAATAACT 618
Db 201 ValValAsnGlnAlaValAsnThrSerAlaProArgMetArgAlaPheSerLeuAlaAla 220
Qy 619 AAAACACTAGATCCCAATAATGGGGGACGCTTTTAGATTCTTAATGAGATTACTACAGCT 678
Db 221 ValAlaAlaAspAlaProAlaAlaGlyThrAspIle---ThrAsnGlnLeuThrAsnVal 239
Qy 679 AATAAGAAATATTAATTAACGTTATCAACTATTAATGAACAA-----AAGACT 726
Db 240 ThrValGlyIleAspSerGlyThrThrValTyProHisGlnAlaGlyTyValLysLeu 259
Qy 727 AAT---GCTGATGCATATCTAATAAGTTTATTATAA-----AAAGTGATTCAA 771
Db 260 AsnTyGlyPheSerValProAsnSerAlaValLysGlyAspThrPheLysIleThrVal 279
Qy 772 AATAATGAACAAAGTTTGTAGGAGCTTTTACAAACGCTTAATGTTTCAACCTTCAAACTAC 831
Db 280 ProLysGluLeuAsnLeuAsnGlyValThrSerThrAlaLysValProPro----- 296
Qy 832 AGTTTGTGCTTTTAGTGTGATGATAACACCGCTCAATTTATAAATATATCAAGAAGGACC 891

```


QY 736 GCATTA-----TCTAATAGTGTATTATAAATA-----GTCATT 768
 Db |||
 811 IleLeuHisSerHisLeuAsnAspThrAsnSerAsnIleArgLysAlaAsnGluIleMet 830
 QY 769 CAAATAATGAACAAAGTTTGTAGGACTTTTACAACGCTAATGTTCAACCTTCAAAAC 828
 Db |||
 831 AsnAsnArgSerGluGluPheLeu-----ArgAsnAla----- 841
 QY 829 TACAGTTTTCTGCTTTTGTAGTGTAGTAAACCCGTCATTTAATAATATGCAAGAAG 888
 Db |||
 842 -----AlaSerGlnAlaGluIleValGlyAlaAsnLysGluArgIleGlnLys 857
 QY 889 ACCGTTTGGATGCT-----GATCAACCTTCAAGTAGAATCTTGCACAAACACG 936
 Db |||
 858 ThrValGluAsnGlySerGlnLeuLeuAspSerLysSerLysAlaIleHisSerAsnSer 877
 QY 937 AATAGTATCACAGATGTTCTGGATTATAGTTTAGCT----- 975
 Db |||
 878 ArgSerMetTyrAsp-----HisCysLeuAlaLeuAlaGluSerGlnLysGln 893
 QY 976 GGAACAAACACGAGTACCAA-----TTTAGT 1002
 Db |||
 894 GlyValAsnLeuGluValGlnThrLeuAspArgLeuLeuGlnLysValLysGluHisSer 913
 QY 1003 TTTAGCAACTPANGCTCCACTCACTGCTTATTTATATTTCCCTATATAAGTTGTTAAAGCA 1062
 Db |||
 914 GluAspAsnThrLysGluLysHisGlnGlnLeu-----LeuAspLeuLeuGluSer 930
 QY 1063 GCTGATGCTAATAACGTTGGATTA-----CAATCAAAATTAATAATGGAATGTTCAA 1116
 Db |||
 931 LeuValGlyAsnAsnAspAsnLeuIleAspSerIleLysThrProHisThrGluLeuGln 950
 QY 1117 CRAAGTTGAG-----TTTGCCACTTCAACTAGTCAAAATAATACTACA----- 1158
 Db |||
 951 LysIleThrAspHisValLeuLysGlyThrThrSerLeuAlaAsnHisThrAsnGluLeu 970
 QY 1159 -----GCTAATCCAACTCCACGAGTTGATGAGATTAAAGTT 1194
 Db |||
 971 LeuGlyLeuGlyAspGluSerLeuCysAsnLeuGluThrThrIleGluAspThrSerLeu 990
 QY 1195 GCTAAATCGTTTATCAGGTTTAAGATTGGCCAAACACAAATC----- 1239
 Db |||
 991 ValLysLeuGluThrThrGlyAspThrProSerLysArgGluLeuProAlaThrProSer 1010
 QY 1240 -----GAATTAAGTGTCCACGGGTGAAGGAATATGAATAAGTTGCG 1284
 Db |||
 1011 TrpThrArgAspSerSerLeuIleLysGluThrThrAsnLeuAsnLeuAspSerAspLys 1030
 QY 1285 CCAATGATGGCAACATTATCTTACCTCAATGAAATAATGCTGATAAGATCCCGGG 1344
 Db |||
 1031 LysPheValArgGluThrThrThrSerSerAsnGlnThrAsnGluProAspVal----- 1048
 QY 1345 TACGCTCGACCC 1356
 Db |||
 1049 TyrAspLysPro 1052

Search completed: October 8, 2003, 18:20:08
 Job time : 57.5855 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:32:42 ; Search time 77.1459 Seconds
(without alignments)
5641.626 Million cell updates/sec

Title: US-09-147-052-1

Perfect score: 2326
Sequence: 1 atgcacatttttagcgaggaa.....gaccocggtacattttataa 1371

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09147052/runat_08102003_154340_29792/app_query.fasta_1.5980
-DB=A_Geneseq_19Jun03 -FORMAT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPGL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052_@CGN_1_1_474_@runat_08102003_154340_29792 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.:

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2324	99.9	456	18	AAW36050
2	2011	86.5	1086	18	AAW36051
3	1910	82.1	615	15	AAW363230
4	1856	79.8	610	15	AAW363229
5	1663	71.5	368	14	AAW44493
6	1663	71.5	368	15	AAW363227
7	1655	71.2	368	16	AAW76955
8	1142	49.1	235	10	AAW36346
9	1142	49.1	235	11	AAW05081
10	1142	49.1	261	11	AAW05082
11	1142	49.1	261	16	AAW79911
12	1117	48.0	261	10	AAW3959
13	807	34.7	661	15	AAW63226
14	807	34.7	661	16	AAW79910
15	715.5	30.8	647	16	AAW11978
16	703	30.2	648	15	AAW56973
17	386.5	16.6	183	10	AAW93649
18	386.5	16.6	183	11	AAW06439
19	314	13.5	865	14	AAW30169
20	284.5	12.2	219	10	AAW93648
21	284.5	12.2	219	11	AAW06438
22	187.5	8.1	1095	22	AAW83030
23	185.5	8.0	10182	23	ABP38314
24	176.5	7.6	6281	22	AAU37403
25	174.5	7.5	933	24	ABU18947
26	174.5	7.5	10498	24	ABU19119
27	174	7.5	2086	22	AAU34143
28	174	7.5	5795	22	AAU37017
29	173.5	7.5	496	23	AAV5490
30	173.5	7.5	933	21	AAV58435
31	173.5	7.5	933	22	AAW69508
32	173.5	7.5	936	18	AAW89801
33	170	7.3	520	23	AAE29262
34	166.5	7.2	2659	23	AAU75489
35	163	7.0	1029	22	AAU34389
36	163	7.0	1048	22	AAU37490
37	161.5	6.9	1237	18	AAW55640
38	161.5	6.9	1237	20	AAU17187
39	160.5	6.9	1072	23	ABB54963
40	158	6.8	2434	22	AAU34339
41	156	6.7	807	21	AAU18311
42	155.5	6.7	2478	22	AAU34320
43	155.5	6.7	2478	22	AAU37374
44	155.5	6.7	2478	24	ABU19002
45	154.5	6.6	682	17	AAW95273

ALIGNMENTS

RESULT 1

AAW36050
ID AAW36050 standard; Protein; 456 AA.

XX

XX AAW36050;

XX 15-JUL-1998 (first entry)

XX Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX Chimeric: Marek's disease virus; outer membrane protein; fusion protein;
antigen; vaccine; poultry.

XX Chimeric - Marek's disease gammaherpesvirus.

XX Chimeric - Mycoplasma gallisepticum.

XX Key Location/Qualifiers

XX Region 1..64

FT /note= "derived from Marek's disease virus gB protein"
 FT 65...456
 XX /note= "derived from M. gallisepticum antigenic protein"
 PN WO9736924-A1.
 XX 09-OCT-1997.
 XX
 PF 28-MAR-1997; 97WO-JP01084.
 XX 29-MAR-1996; 96JP-0103548.
 PR (JAPG) NIPPON ZEON KK.
 XX Saito S, Tsuzaki Y, Yanagida N;
 XX WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 456 AA;
 Alignment Scores:
 Pred. No.: 2,25e-191 Length: 456
 Score: 2324.00 Matches: 456
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 18 Gaps: 0
 US-09-147-052-1 (1-1371) x AAW36050 (1-456)
 QY 1 ATGCACTATTTAGCGGAATGTCATATTTTCCTATTAGTTATTTCTATATGGTACGAC 60
 DB 1 MethIstYrPheArgAsnCysIlePhePheLeuIleValIleLeuIleGlyThrAsn 20
 QY 61 TCATCTCGAGTACCCAAATGTGACATCAAGAGAGAGTGTTCGAGCGTCCAGTTGTCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGCTCTACGTTTATCTTTGTCGCCACCGAGTGGGTTCACCGTGTATCGGTCTA 180
 DB 41 GluGluGluSerThrPheThrLeuGluSerProProValGlySerThrValIleArgLeu 60
 QY 181 GAATTCGGCTCTATCTCTATTACTAAAAAGATGCAAAACCCCAATTAATGGCCAAACCCAA 240
 DB 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
 QY 241 TTAGAAGCAGCGGATGGAGTTACAGATCTAATCAATGCTAAAGCGATGACATTAGCT 300
 DB 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
 QY 301 TCATCTACAGACTATGCCAAGATTCAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 360
 DB 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
 QY 361 ACAGTTACATAACCTTTAATGCAACATPAGAACAACCTAAAAATGGCTAAACCTAATTTA 420
 DB 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
 QY 421 GAATCAGCCATCAACCAAGCTAATACGGATTAACGACTTTTGATATGAACACCCAAAT 480

Db 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrPheAspAsnGluHisProAsn 160
 QY 481 TTAGTTGAAGCATACAAAGCAGCTAAACCCTTTAGAACAAAGCTCTACTAACCTTGAA 540
 DB 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
 QY 541 GGTTCGTCATCAACTGCTTATTAATCAATTCGCAATATTTAGTGGATCTATACATAAAA 600
 DB 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
 QY 601 GCTAGTAGTTTAATACTAAACACACTAGATCCACTAAATGGGGGAGCCTTTTACATCTCT 560
 DB 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
 QY 661 AATGAGATTACTACAGCTAATAAGAAATATTAATATACGTTATCACTATTAATCAACAA 720
 DB 221 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
 QY 721 AAGACTAATGCTGATGCATTATCTAATAGTTTATTAATAAAGTAGTTCAAAATTAATGAA 780
 DB 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGlu 260
 QY 781 CAAAGTTTGTAGGCACTTTTACAAACGCTAATGTTCACCTTCAACCTTCAACCTTGTGT 840
 DB 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280
 QY 841 GCTTTTGTAGTGTGATGTAACACCCGCTCAATTAATTAATATGCAAGAGACCGTTGGAA 900
 DB 281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn 300
 QY 901 GGTGATGAACCTTCAAGTAGAATTTCTGCAACACGAATAGTATACACAGATGTTCTTGG 960
 DB 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTirp 320
 QY 961 ATTTATAGTTTGTAGTGGAAACACGAGTACCAATTTAGTTTGTAGCAACTATGGTCCA 1020
 DB 321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
 QY 1021 TCAACTGGTTATTTATTTTCCCTTATTAAGTTGGTTAAAGCAGCTGATCTAATAACGTT 1080
 DB 341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
 QY 1081 GGATTACATACAAATAATAATGGAATGTTCAACAAGTTGAGTTGGCAGCTTCAACT 1140
 DB 361 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
 QY 1141 AGTCAATAATACTACAGCTAATCCCACTCCAGCTGATGATGAGATTAAAGTTGCTAAA 1200
 DB 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
 QY 1201 ATCGTTTATCAGTTTAAAGATTTGGCCAAACACAATCGAATTAAGTTGTTCCAAACGGT 1260
 DB 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
 QY 1261 GAAGGAATATGAATAAGTTGCCCAATGATTGGCAACATTTATCTTAGCTCAAAATGAA 1320
 DB 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
 QY 1321 AATAATGCTGATAGATCCCGGTACCGTACCGTACCGTACCGTACCGTACCGTACCGT 1368
 DB 441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

RESULT 2

AAW36051

ID AAW36051 standard; Protein; 1086 AA.

XX

AC AAW36051;

XX

DT 15-JUL-1998 (first entry)

XX

DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.

XX

KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 XX antigen; vaccine; poultry.
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 FH Key Location/Qualifiers
 FT Region 1..672
 FT /note= "derived from Marek's disease virus gB protein"
 FT Region 693..1086
 FT /note= "derived from M. gallisepticum antigen"
 FT W09736924-Al.
 PN PN
 XX 09-OCT-1997.
 PD
 XX 28-MAR-1997; 97WO-JP01084.
 XX 29-MAR-1996; 96JP-0103548.
 XX (JAPG) NIPPON ZBON KK.
 PA Saito S, Tsuzaki Y, Yanagida N;
 PI WPI: 1997-503046/46.
 DR N-PSDB; AAT96396.
 XX
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;
 Alignment Scores:
 Pred. No.: 2.17e-164 Length: 1086
 Score: 2011.00 Matches: 395
 Percent Similarity: 99.75% Conservative: 0
 Best Local Similarity: 99.75% Mismatches: 1
 Query Match: 86.46% Indels: 0
 DB: 18 Gaps: 0
 US-09-147-052-1 (1-1371) x AAW36051 (1-1086)
 QY 181 GAATTCGGCTGTATCTCTATTACTAAAAAGATGCAACCCCAATTAATGGCCAAACCCAA 240
 DB 691 GluPheGlyCysMetSerIleThrLysLysPalaAsnProAsnAsnGlyGlnThrGln 710
 QY 241 TTAGAACGACGCGCAATGGAGTAAACAGATCTCAATCAATGCTAAACGCGATGACATTAGCT 300
 DB 711 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 730
 QY 301 TCACACAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTTATAGTGAACCTGAA 360
 DB 731 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 750
 QY 361 ACAGTTAAACAATCACTTAAATGCAACATTAGAACAACTAAATAATGGCTAAACACTAATTTA 420
 DB 751 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 770
 QY 421 GAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGTATTAATCAACACCCAAAT 480
 DB 771 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 790
 QY 481 TTAGTTGAGCATACAAAGCACTAAACACCACCTTTAGAACAAACGCTGCTACTAACCTTGA 540

Db 791 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 810
 QY 541 GCTTTTGTCACTCAACTGCTTATTAATCAAAATTCCTCAATAATTTAGTGGATCTATACATAAA 600
 DB 811 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 830
 QY 601 GCTAGTAGTTTAATTAACATAACACATAGATCCACATAAATGGGGGACGCTTTTATAGTCT 660
 DB 831 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 850
 QY 661 AATGAGATTACTACAGCTAATTAAGATATTAATAATACCTTTATCACTATTATTAACACAA 720
 DB 851 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 870
 QY 721 AAGACTAATGCTGATGCTATTAATAGTTTATTAATAAAGTATTCAAAATTAATGAA 780
 DB 871 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 890
 QY 781 CAAGCTTTGTAGGACCTTTTACAAACGCTAATGTTCAACCTTCAACATACAGTTTGT 840
 DB 891 GlnGlyPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 910
 QY 841 GCTTTTGTAGTGTATGATACACCCCTCAATTAATAATATGCAAGAGACCTTTGGAT 900
 DB 911 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTrpAsn 930
 QY 901 GGTGATGAACCTTCAAGTAGAATCTTGCAACACAGATAGTATCACAGATCTTCTCTGG 960
 DB 931 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 950
 QY 961 ATTTATAGTTTGTAGTGGAAACAAACAGAGTACCAATTTAGTTTGAACACTATGCTCA 1020
 DB 951 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 970
 QY 1021 TCACTGCTTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATACGTT 1080
 DB 971 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 990
 QY 1081 GGATTACAAATCAAAATTAATAATGAATGTTCAACAGTTGAGTTTGCCTCACTTCAACT 1140
 DB 991 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 1010
 QY 1141 AGTCAAAATATATCTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTCTGCTAA 1200
 DB 1011 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 1030
 QY 1201 ATCGTTTATCAGGTTTAAAGATTGCGCCAAACACAACTGAATTAAGTGTCCACGCGT 1260
 DB 1031 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 1050
 QY 1261 GAAGCAATATCAATTAAGTTGCGCAATGATTGCAACATTTATCTTACCTCAAAATGAA 1320
 DB 1051 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 1070
 QY 1321 AATAATGCTGATPAAGATCCCGGGTACCGTCCGACCGGCTACATTTTTA 1368
 DB 1071 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 1086
 RESULT 3
 AAR63230
 ID AAR63230 standard; Protein; 615 AA.
 XX
 AC AAR63230;
 XX
 DT 25-MAR-2003 (updated)
 DT 23-JUN-1995 (first entry)
 XX
 DE Mycoplasma gallisepticum antigen (UM-67).
 XX
 KW recombinant avipox virus; live vaccine; mycoplasma antigen.
 XX
 OS Mycoplasma gallisepticum.

```

XX FH      Location/Qualifiers
FT Protein 1..615
FT /note= "Trip residues correspond to TGA codons"
XX
XX PN      W09423019-A1.
XX PD      13-OCT-1994.
XX
XX PF      31-MAR-1994; 94WO-JP00541.
XX
XX PR      31-MAR-1993; 93JP-0074139.
XX PR      30-SEP-1993; 93JP-0245625.
XX
XX PA      (JAPG ) NIPPON ZEON KK.
XX PA      (SHIO ) SHIONOGI & CO LTD.
XX
XX PI      Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX PI      Saeiki S, Saïtos, Takahashi K;
XX
XX DR      WPI; 1994-333181/41.
XX DR      N-PSDB; AAQ77857.
XX
XX PT      Recombinant avipox virus combining DNA encoding a polypeptide -
XX PT      exhibiting antigenicity of mycoplasma, useful for the production
XX PT      of a live vaccine
XX
XX PS      Claim 4; Page 87-91; 123pp; Japanese.
XX
XX CC      A restriction fragment of the insert of M.gallisepticum genomic clone
XX CC      PUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX CC      ORF encodes an antigenic polypeptide (AAR63230). A recombinant avipox
XX CC      virus comprising the coding sequence can be used as a live vaccine to
XX CC      protect against infection by Mycoplasma gallisepticum.
XX CC      (Updated on 25-MAR-2003 to correct FN field.)
XX
XX SQ      Sequence 615 AA;

Alignment Scores:
Pred. No.: 9.22e-156 Length: 615
Score: 1910.00 Matches: 376
Percent Similarity: 98.96% Conservativity: 3
Best Local Similarity: 98.17% Mismatches: 4
Query Match: 82.12% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-1 (1-1371) x AAR63230 (1-615)

QY 190 TGTATGCTATTACTAAAAAGATGCAACCCCAAAATAATGGCCAAACCCCAATTAGAAGCA 249
DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
QY 250 GCGGAATGGAGTTAAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCCACTACAA 309
DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
QY 310 GACTATGCCAAGATTGAAGCTAGTTATCATCGCTTATAGTAGAAGCTGAACAGTTAAC 369
DB 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
QY 370 AATAACCTTAATGCAACATTAGAACAACTAAATGGCTAAACACTAATTAGAATCAGCC 429
DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGGATAAAGCACTTTTGATAATGAACACCAAAATTTAGTTGAA 489
DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
QY 490 GCATACAAAGCACTAAAACCACTTTAGAACACGCTGCTACTAAGCTTGAAGTTGTGCA 549
DB 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
QY 550 TCAACTGCTTATAATCAAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGT 609

```

```

|||||
147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
QY 610 TTATACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCTTAATGAGATT 669
DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
QY 670 ACTACAGCTAATAAGAATATTAATAACGTTATCAACTATTATTAATGAACAAAGACTAAT 729
DB 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GCTGATGCATTATCTAATAGTTTATTAATAAAGCTGATTCAAAATAATGAACAAAGTTT 789
DB 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226
QY 790 GTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTAGT 849
DB 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
QY 850 GCTGATGTAACACCCGTCGAATTATTAATATGCAAGAGGACCGTTTGGATGCTGATGAA 909
DB 247 AlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValIleAsnGlyAspGlu 266
QY 910 CCTTCAAGTAGAATTTCTGCAACACAGATAGTATCACAGATGTTCTTGGATTATAGT 969
DB 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerThrIleTyrSer 286
QY 970 TTAGCTGGAACAAACACGAGTAGTACCAATTTAGTTTAGCAACTATGTCCTCACTGGT 1029
DB 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
QY 1030 TATTTATTTCCCTTATTAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAA 1089
DB 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGln 326
QY 1090 TACAAATTAATATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAAAT 1149
DB 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
QY 1150 AATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATTCGTTTA 1209
DB 347 AsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLysIleValLeu 366
QY 1210 TCAGGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTTCCACGGGTGAAGGAAAT 1269
DB 367 SerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGlyGluGlyAsn 386
QY 1270 ATCAATAAAGTTGCGCAATGATTTGCAACATTTATCTTAGCTCAATCAAAATATAGCT 1329
DB 387 MetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGluAsnAla 406
QY 1330 GATAGATC 1338
DB 407 AspLysIle 409

RESULT 4
AAR63229
ID AAR63229 standard; Protein; 610 AA.
XX
AC AAR63229;
XX
DT 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
XX
DE Mycoplasma gallisepticum antigen (UM-66).
XX
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Protein 1..610
FT /note= "Trip residues correspond to TGA codons"

```

XX WO9423019-A1.
 PN
 XX
 PD 13-OCT-1994.
 XX
 PF 31-MAR-1994; 94WO-JP00541.
 XX
 PR 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX
 PA (JAP) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeki S, Saitos, Takahashi K;
 XX
 DR N-PSDB; AAQ77856.
 DR
 XX
 PT Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 PT of a live vaccine
 XX
 PS Claim 4; Page 78-81; 123pp; Japanese.
 XX
 CC A restriction fragment of the insert of M.gallisepticum genomic clone
 CC PUM-66 containing an open reading frame was sequenced (AAQ77856). The
 CC ORF encodes an antigenic polypeptide (AAQ63229). A recombinant avipox
 CC virus comprising the coding sequence can be used as a live vaccine to
 CC protect against infection by Mycoplasma gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 610 AA:
 Alignment Scores:
 Pred. No.: 4, 03e-151 Length: 610
 Score: 1856.00 Matches: 365
 Percent Similarity: 58.17% Conservative: 10
 Best Local Similarity: 99.55% Mismatches: 7
 Query Match: 79.79% Indels: 0
 DB: 15 Gaps: 0
 US-09-147-052-1 (1-1371) x AAQ63229 (1-610)
 QY 190 TGTATGCTTACTATAAAGATGCAACCCAAATAATAGCCAAACCAATTAGAACA 249
 DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGluAla 46
 QY 250 GCGCGAATGGAGTTAACAGATCTAATCAATGCTTAAGCGATCACATTAGCTTCACTACAA 309
 DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGln 66
 QY 310 GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGCTTAAC 369
 DB 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValasn 86
 QY 370 AATAACCTTAATGCAACATTAGAACACTAAATAAATGGCTAAACATAATTAGATCAGCC 429
 DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 QY 430 ATCAACCAAGCTTAATACCGATAAAGAGCTTTTGTATGATGATGAACCCCAATTTAGTTGAA 489
 DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
 QY 490 GCATAAAGCAAGCACTAAACCACTTTAGAACACGCTGCTACTAACCTTGAAGGTTGTGCA 549
 DB 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSer 146
 QY 550 TCAACTGCTTATTAATCAAAATTCGCAATATTTAGTGGATCTATACAATAAAGCTAGTAGT 609
 DB 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSer 166
 QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTATAGATTCTTAATGAGATT 669

Db 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIle 186
 QY 670 ACTACAGCTAATAAGAATATAATAATAGCTTATCACTATTAATGAACAAAGACTAAT 729
 Db 187 ThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
 QY 730 GCTGATGCATTATCAATAGTTTATTAATAAAGTGAATTCAAAATAATGAACAAAGTTTT 789
 Db 207 AlaAspAlaLeuAlaAsnSerPheIleLysGluValIleGlnAsnAsnLysGlnSerPhe 226
 QY 790 GTAGGACTTTTACAAACGCTTAATGTTCAACCTTCAACTACAGTTTGTCTCTTTAGT 849
 Db 227 ValGlyMetPheThrAsnThrAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
 QY 850 GCTGATGTAACACCGCTCAATTAATAATATGCAAGGACCGTTTGAATGGTGGATGAA 909
 Db 247 AlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTrpAsnGlyAspGlu 266
 QY 910 CCTTCAAGTAGAATTCCTTCAACACAGATAGTATCACAGATGTTTCTTGGATTATAGT 969
 Db 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrpIleTyrSer 286
 QY 970 TTAGCTGGAACAAACACGAGTACCAATTTAGTTTACCACTATGCTCCACTCACTGGT 1029
 Db 287 LeuSerGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
 QY 1030 TATTTATATTCCTTATAAGTTGGTTAAAGCAGCTGATGCTTAATAACGTTGGATTACAA 1089
 Db 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaSerAsnValGlyLeuGln 326
 QY 1090 TACAAATTAATAATGGAATGTTTCAACAAGTTGAGTTGGCAGTTCACTAGTGCAGAT 1149
 Db 327 TyrLysLeuAsnAsnGlyAsnValGlnProValGluPheAlaThrSerThrSerAlaAsn 346
 QY 1150 AATACTACAGCTAATCCAACTCCAGCTTGATGATAGATTAAAGTTGCTAAAATCGTTTA 1209
 Db 347 AsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLysIleValLeu 366
 QY 1210 TCAGGTTTAAGATTGGCCAAACCAATCGAATTAAGTTGTTCCACGCGTGAAGGAAT 1269
 Db 367 SerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGlyGluArgAsn 386
 QY 1270 ATGAATAAGTTGGCCCAATGATGCGCAACATTATCTTAGCTCAAAATGAATAATGCT 1329
 Db 387 MetAsnLysValAlaProMetIleGlyAsnMetTyrIleThrSerSerAsnAlaGluAla 406
 QY 1330 GATAAG 1335
 Db 407 AsnLys 408
 RESULT 5
 AAR44493
 ID AAR44493 standard; Protein; 368 AA.
 XX AAR44493;
 AC AAR44493;
 XX 25-MAR-2003 (updated)
 DT 16-JUN-1994 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kD antigen.
 XX
 KW Vaccine; mycoplasma infection; poultry; fowl.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262
 FT /note= "corresponds to NNN codon in AAQ53419"
 FT Misc-difference 283
 FT /note= "corresponds to NNN codon in AAQ53419"
 XX
 PN W09324646-A1.

XX 09-DEC-1993.
 XX 28-MAY-1993; 93WO-JP00715.
 XX 29-MAY-1992; 92JP-0138819.
 XX (JAPC) NIPPON ZEON KK.
 XX (SHIO) SHIONOGI & CO LTD.
 XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
 XX WPI; 1993-405837/50.
 XX N-PSDB; AAQ53419.
 XX Mycoplasma gallisepticum antigen and DNA coding for it - useful
 XX for vaccination of fowl against mycoplasma infections
 XX Claim 2; Page 23-26; 37pp; Japanese.
 XX The sequence coding for the 40kDa antigen was obtained by PCR
 XX amplification of M.gallisepticum genomic DNA. The antigen reacts
 XX with Mycoplasma-immune or Mycoplasma-infected serum and can be used
 XX as a vaccine to protect fowl from M.gallisepticum infection.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 368 AA;
 Alignment Scores:
 Pred. No.: 1,41e-134 Length: 368
 Score: 1663.00 Matches: 333
 Percent Similarity: 97.96% Conservative: 3
 Best Local Similarity: 97.08% Mismatches: 6
 Query Match: 71.50% Indels: 1
 DB: 14 Gaps: 0
 US-09-147-052-1 (1-1371) x AAR44493 (1-368)
 QY 190 TGTATGCTATTACTATAAAGAGTGCACCAACCAATAATGCGCAACCCCAATAGAGCA 249
 Db 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
 QY 250 GCGCGAATGAGTAAACAGATCTAATCAATGCTAAACGCGATGACATTAGCTTCACTACAA 309
 Db 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
 QY 310 GACTATGCCAAGATGAGCTAGTCTTATCATCTGCTTATAGTGAAGCTGAACAGTTAAC 369
 Db 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
 QY 370 AATAACCTTAATGCACATTAGAACCACTAAATAATGGCTAAACCTAATTTAGAACTAGCC 429
 Db 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 QY 430 ATCAACCAAGCTAATACGGATAAACACGCTTTTGATAATGAACACCCCAATTTAGTTGAA 489
 Db 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
 QY 490 GCATACAAAGCACTAAACCACTTTAGAACACGCTGCTACTAACCTTGAAGTTTGTCAC 549
 Db 127 AlaTyrIleAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
 QY 550 TCAACTGCTTATATCAATTCGCAATTAATTTAGTGGATCTATACAAATAAGCTAGTACT 609
 Db 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
 QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGAAACGCTTTTATAGATTTCTAATGAGATT 669
 Db 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
 QY 670 ACTACAGCTAATAAGAAATTAATTAATAGCTTATCACTATTAAATCAACAAAGACTAAT 729
 Db 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206

QY 730 GCTGATGCAATCTAATAGTATTTTATAAATAAGTATCAAAATAATGAACAAAGTTT 789
 Db 207 AlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGluGlnSerPhe 226
 QY 790 GTAGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGT 849
 Db 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
 QY 850 GCTGATGTACACCCGCTCAATTATTAATATGCAAGAAGACCGCTTGAATGCTGATGAA 909
 Db 247 AlaAspValThrProValAsnTyrLysTyrAlaArgThrVal***AsnGlyAspGlu 266
 QY 910 CTTTCAAGTAGAATCTTGCACCAACCAATAGTATACAGATGTTCTTGGATTATAGT 969
 Db 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSer***IleTyrSer 286
 QY 970 TTAGCTGGAACAACACGAGTACCAATTTAGTTTGTAGCAACTATGCTCCATCAACTGGT 1029
 Db 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
 QY 1030 TATTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATACGTTGGATTACAA 1089
 Db 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGln 326
 QY 1090 TACAAATTAATAATGCAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAAT 1149
 Db 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
 QY 1150 AATACTACAGCTAATCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAATCGTTTAA 1209
 Db 347 AsnThrThrAlaAsnProThr-GlnGlnLeuMetArgLeuLysLeuLysSerPheTy 366
 QY 1210 TCAGGTT 1216
 Db 366 rGlnVal 368
 RESULT 6
 AAR63227
 ID AAR63227 standard; Protein; 368 AA.
 XX AAR63227;
 AC XX
 DT 25-MAR-2003 (updated)
 DT 23-JUN-1995 (first entry)
 XX Mycoplasma gallisepticum 40kD antigen.
 DE recombinant avipox virus; live vaccine; mycoplasma 40kD antigen;
 KW TTM-1.
 KW Mycoplasma gallisepticum.
 OS Mycoplasma gallisepticum.
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to a NNN codon"
 FT Misc-difference 283 /note= "corresponds to a NNN codon"
 FT W09423019-A1.
 PN 13-OCT-1994.
 PD 31-MAR-1994; 94WO-JP00541.
 PF 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX (JAPC) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeki S, Saitos, Takahashi K;

XX WPI: 1994-333181/41.
 DR N-PSDB; AAQ77854.
 XX Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 of a live vaccine
 XX
 PS Claim 4; Page 71-74; 123pp; Japanese.
 XX The plasmid pUTM-1p contains a sequence (the TTM-1 gene) coding
 CC for the 40kD antigen of Mycoplasma gallisepticum under the control
 CC of a synthetic promoter. A 1300 bp restriction fragment containing
 CC the promoter-ORF sequence was excised and was used in the
 CC construction of plasmid pN7929-R2. This in turn was involved in the
 CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FpV
 CC sequences. The recombinant avipox virus is useful as a live vaccine
 CC to protect against infection by Mycoplasma gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 368 AA;
 Alignment Scores:
 Pred. No.: 1,41e-134 Length: 368
 Score: 1663.00 Matches: 333
 Percent Similarity: 97.96% Conservative: 3
 Best Local Similarity: 97.08% Mismatches: 6
 Query Match: 71.50% Indels: 1
 DB: 15 Gaps: 0
 US-09-147-052-1 (1-1371) x AAR63227 (1-368)
 QY 130 TGTATGCTATTACTAAAGATGCAACACCAATAATGGCCAAACCAATTTAGAAGCA 249
 DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
 QY 250 GCGCAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCTTCACATAA 309
 DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
 QY 310 GACTATGCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 369
 DB 67 AspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGluThrValAsn 86
 QY 370 AATAACCTTAATGCAACATTAAGAACCACTAAATAATGGCTAAACCTTAATTTAGAAATCAGCC 429
 DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 QY 430 ATCAACCAAGCTAATACGGATAAAACGACTTTTGATATGAACACACCAATTTAGTTGAA 489
 DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 126
 QY 490 GCATACAAAGCACTAAACCACTTTAGAACACCTGCTACTTAACCTTGAAGGTTGTCA 549
 DB 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
 QY 550 TCAACTGCTTATTAATCAAAATGCGCAATAATTTAGTGGATCTATACAAATAAAGCTAGTAGT 609
 DB 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
 QY 610 TTAATACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTTAATGAGATT 669
 DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
 QY 670 ACTACAGCTAATAAGATATTAATAATACCTTATCACTATTATTAATGAACAAAGACTAAT 729
 DB 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
 QY 730 GCTGATGCTATCATCATAGTTTATTAATAAAGTGATTCAAAATAATGAACAAAGTTT 789
 DB 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226

QY 790 GTAGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTGCTTTAGT 849
 DB 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
 QY 850 GCTGATGTAACACCGCTCAATTTATAAATATGCAACGACGCGTTTGGAAATGTCATGAA 909
 DB 247 AlaAspValThrProValAsnTyrLysTyrAlaArgThrVal**AsnGlyAspGlu 266
 QY 910 CCTCAAGTAGAATTTCTTGCACAAACGAATAGTATCATCAGATGTTCTTGGATTTATAGT 969
 DB 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSer***IleTyrSer 286
 QY 970 TTAGCTGGAACAAACACGAAAGTACCAATTTAGTTTGTAGCACTATGTCCTCACTCACTGGT 1029
 DB 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
 QY 1030 TATTTATATTTCCCTTATTAAGTTGCTTAAAGCAGCTGATGCTAATAACGTTGCGATTACAA 1089
 DB 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGln 326
 QY 1090 TACAAATTAATAATGGAATGTTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAT 1149
 DB 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
 QY 1150 AATACTACAGCTAATCCCACTCCAGCAGTTGATGATGAGATTAAAGTTGCTAAATCGTTTAA 1209
 DB 347 AsnThrThrAlaAsnProThr-GlnGlnLeuMetArgLeuLysLeuLeuLysSerPheTyr 366
 QY 1210 TCAGGTT 1216
 DB 366 rGlnVal 368
 RESULT 7
 AAR76955
 ID AAR76955 standard; Protein; 368 AA.
 XX
 AC AAR76955;
 XX
 DT 25-MAR-2003 (updated)
 DT 12-MAR-1996 (first entry)
 DE Mycoplasma gallisepticum antigenic protein TTM-1.
 XX
 KW Antigenic protein; vaccine; poultry; diagnosis; TTM-1.
 OS Mycoplasma gallisepticum.
 FH Key Location/Qualifiers
 FT Misc-difference 282 /note= "any amino acid"
 FT Misc-difference 283 /note= "any amino acid"
 FT
 XX JP07133295-A.
 XX
 PD 23-MAY-1995.
 XX
 PF 27-AUG-1993; 93JP-0213102.
 XX
 PR 27-AUG-1993; 93JP-0213102.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 DR WPI: 1995-220782/29.
 DR N-PSDB; AAQ94711.
 XX
 PT A new antigenic protein which reacts with Mycoplasma gallisepticum -
 PT is useful in a component vaccine for use against poultry infected
 PT with M. gallisepticum.
 XX
 PS Claim 6; Figs 5-6; 33pp; Japanese.

```

XX AA094711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
CC infectious diseases in poultry, and as a diagnostic agent for
CC M. gallisepticum.
XX (Updated on 25-MAR-2003 to correct PA field.)
SQ Sequence 368 AA;

Alignment Scores:
Pred. No.: 6.85e-134 Length: 368
Score: 1655.00 Matches: 331
Percent Similarity: 97.67% Conservative: 4
Best Local Similarity: 96.50% Mismatches: 7
Query Match: 71.15% Indels: 1
DB: 16 Gaps: 0

US-09-147-052-1 (1-1371) x AAR76955 (1-368)
QY 190 TGTATGCTATTAATAAAGATGCAACCCCAAAATATGCGCAACCCCAATAGAACGA 249
DB 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
QY 250 GCGCAATGAGGTAAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAA 309
DB 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66
QY 310 GACTATGCCAAGATTGAGCTAGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAAC 369
DB 67 AspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
QY 370 AATAACCTTAATGCAACATTAGAACCACTAAATAATGCTAAACACTTAATTTAGATCAGCC 429
DB 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuIleAsnAlaLysAlaSerLeuGluSerAla 106
QY 430 ATCAACCAAGCTAATACGGATAAAGACACTTTTGTATAATGAACACCAATTTAGTTGAA 489
DB 107 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGln 126
QY 490 GCATACAAAGCACTAAACACCTTTAGACACAGCTGCTACTAACCCTTGAAGTTTGTCA 549
DB 127 ProTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
QY 550 TCAACTGCTTATTAATCAAAATTCGCAATTAATTTAGTGGATCTATACAAATAAGCTACTAGT 609
DB 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
QY 610 TTAATACTAAACACTAGATCCACTAAATGCGGGAACGCTTTTAGATTCTAATGAGATT 669
DB 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
QY 670 ACTACAGCTAATAAGAATATTAATAATAGTTATCACTTATCACTATTAAATGAACAAAGCTAAT 729
DB 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
QY 730 GGTGATGCATTAATTAATAGTTTATTAATAAAGTGATTCAAAATAATGAACAAAGCTTTT 789
DB 207 AlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGluGlnSerPhe 226
QY 790 GTAGGACCTTTTACAACCTTAATGTTCAACCTTCAAACTACAGTTTGTCTTTTGTAGT 849
DB 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
QY 850 GCTGATGTAAACCCGCTCAATTAATAATGAAGAGGACCGTTTGGNATGCTGATGAA 909
DB 247 AlaAspValThrProValAsnTyrLysTyrAlaArgThrVal***AsnGlyAspGlu 266
QY 910 CCTCAAGTAGAATTTCTTGAACACGATAGTATCATCAGATGTTTCTTGATTTATAGT 969
DB 267 ProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSer***IleTyrSer 286
QY 970 TTAGCTGGACAAACACCAAGTACCAATTTAGTTTACCACTATGGTCCATCAACTGGT 1029

```

```

DB 287 LeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGly 306
QY 1030 TATTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAA 1089
DB 307 TyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGln 326
QY 1090 TACAAATTAATTAATGAAATGTTCAACAAGTTTCAGTTTCCACTTCAACTAGTGCAGAT 1149
DB 327 TyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaAsn 346
QY 1150 AATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAATCGTTTAA 1209
DB 347 AsnThrThrAlaAsnProThr-GlnGlnLeuMetArgLeuLysLeuLysSerPheTy 366
QY 1210 TCAGGTT 1216
DB 366 rGlnVal 368

```

```

RESULT 8
AAP93646
ID AAP93646 standard; protein; 235 AA.
XX
AC AAP93646;
XX
DT 25-MAR-2003 (updated)
DT 11-MAY-1990 (first entry)
XX

```

Amino acid sequence of Mycoplasma gallisepticum (MGI) polypeptide.

Mycoplasma gallisepticum; Poultry vaccine; ss;

Mycoplasma gallisepticum.

EP345021-A.

06-DEC-1989.

31-MAY-1989; 89EP-0305441.

31-MAY-1989; 89EP-0305441.

02-JUN-1988; 88JP-0136343.

(JAPG) NIPPON ZEON KK.

(SHIO) SHIONOGI SEIYAKU KK.

(JAPG) NIPPON ZEON KK.

(SHIO) SHIONOGI SEIYAKU KK.

Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

WPI; 1989-358933/49.

N-PSDB; AAN92568.

New antigenic proteins of Mycoplasma gallisepticum - useful as poultry vaccines.

Disclosure; Fig.1a; 3lpp; English.

This amino acid sequence of MGI is encoded by M1 DNA and elicits an antigen-antibody reaction with anti-MG poultry sera. It can be used as a vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 235 AA;

Alignment Scores:

Pred. No.: 7.75e-90 Length: 235
Score: 1142.00 Matches: 228
Percent Similarity: 98.30% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 4

```
Query Match: 49.10% Indels: 0
Db: 10 Gaps: 0
US-09-147-052-1 (1-1371) x AAP93646 (1-235)
QY 190 TGTATGCTATTACTAAAAAGATGCAAAACCAATATAGCCAAACCAATTAGAAGCA 249
DB 1 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 20
QY 250 GCGGAATGGAGTAAACAGATCTAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAA 309
DB 21 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 40
QY 310 GACTATGCCAAGATGAGCTAGTTATCATCTGCTTATAGCTGAGCTGAACAGCTTAAAC 369
DB 41 AspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGluThrValAsn 60
QY 370 AATAACCTTAATGCAACATTAAGAACACTAAATGCTAAACCTTAATGCTTCACTACAA 429
DB 61 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 80
QY 430 ATCAACCAAGCTTAATACGGATAAAGACTTTTGTGATAATGAACACCAATTTAGTTGAA 489
DB 81 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 100
QY 490 GCATACCAAGCTTAATGCAACCTTAATGCAACCTTAATGCAACCTTAATGCAACCTTA 549
DB 101 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 120
QY 550 TCAACTGCTTATTAATCAAAATTCGCAATATTTAGTGGATCTATACAAATAAAGCTAGT 609
DB 121 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 140
QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGAACGCTTTTATAGATCTCTATAGAT 669
DB 141 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 160
QY 670 ACTACAGCTTAATGAATATTAATGCTTATCACTTATCACTTATCACTTATCACTTAT 729
DB 161 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 180
QY 730 GCTGATGCTATCATTAATGATTTTATTAATAAAGTGAATCAAAATAATGAACAAAGTTT 789
DB 181 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 200
QY 790 GTAGGACTTTTACAAACGCTTAATGCTTCACTTCACTTCACTTCACTTCACTTCACT 849
DB 201 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 220
QY 850 GCTGATGTAACACCGCTCAATTAATAATATCAAGAGGACCGTT 894
DB 221 AlaAspValThrProValAsnTyrLysTyrAlaArgThrVal 235
RESULT 9
AAR05081
ID AAR05081 standard; protein; 235 AA.
XX
AC AAR05081;
XX
DT 10-MAR-2003 (updated)
DT 08-OCT-1990 (first entry)
XX
DE MG-1 antigen.
XX
KW Mycoplasma gallisepticum; poultry; vaccine.
XX
OS Mycoplasma gallisepticum.
XX
PN JP02111795-A.
XX
PD 24-APR-1990.
XX
PF 02-JUN-1989; 89JP-0136343.
XX
PR 02-JUN-1989; 89JP-0136343.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI KK.
XX
DR WPI; 1990-169109/22.
DR N-PSDB; AAQ04686.
XX
PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
PT antigen protein of the disease and recombinant vector
PT incorporated with its coding gene.
XX
PS Claim 2; Fig 1a; 20pp; Japanese.
XX
CC DNA encoding the protein can be inserted into an expression vector
CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
CC be ligated to other DNA to produce fusion proteins with an N-terminal
CC bacterial enzyme sequence.
CC See also AAR05081-2 and AAR06437-41.
CC (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ Sequence 235 AA:
Alignment Scores: 7.75e-90 Length: 235
Pred. No.: 1142.00 Matches: 228
Score: 98.30% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 4
Query Match: 49.10% Indels: 0
DB: 11 Gaps: 0
US-09-147-052-1 (1-1371) x AAR05081 (1-235)
QY 190 TGTATGCTATTACTAAAAAGATGCAAAACCAATATAGCCAAACCAATTAGAAGCA 249
DB 1 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 20
QY 250 GCGGAATGGAGTAAACAGATCTAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAA 309
DB 21 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 40
QY 310 GACTATGCCAAGATGAGCTAGTTATCATCTGCTTATAGCTGAGCTGAACAGCTTAAAC 369
DB 41 AspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGluThrValAsn 60
QY 370 AATAACCTTAATGCAACATTAAGAACACTAAATGCTAAACCTTAATGCTTCACTACAA 429
DB 61 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 80
QY 430 ATCAACCAAGCTTAATACGGATAAAGACTTTTGTGATAATGAACACCAATTTAGTTGAA 489
DB 81 IleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGlu 100
QY 490 GCATACCAAGCTTAATGCAACCTTAATGCAACCTTAATGCAACCTTAATGCAACCTTA 549
DB 101 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 120
QY 550 TCAACTGCTTATTAATCAAAATTCGCAATATTTAGTGGATCTATACAAATAAAGCTAGT 609
DB 121 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 140
QY 610 TTAATAACTAAACACTAGATCCACTAAATGGGGAACGCTTTTATAGATCTCTATAGAT 669
DB 141 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 160
QY 670 ACTACAGCTTAATGAATATTAATGCTTATCACTTATCACTTATCACTTATCACTTAT 729
DB 161 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 180
QY 730 GCTGATGCTATCATTAATGATTTTATTAATAAAGTGAATCAAAATAATGAACAAAGTTT 789
DB 181 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 200
QY 790 GTAGGACTTTTACAAACGCTTAATGCTTCACTTCACTTCACTTCACTTCACTTCACT 849
DB 201 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 220
QY 850 GCTGATGTAACACCGCTCAATTAATAATATCAAGAGGACCGTT 894
DB 221 AlaAspValThrProValAsnTyrLysTyrAlaArgThrVal 235
```

Db 181 AlaaspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlnInSerPhe 200
 QY 790 GTAGGAGCTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGCTTTGTTGCTTTAGT 849
 Db 201 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 220
 QY 850 GCTGATGTAACACCGCTCAATTAATAATATGCAAGAGGACCGTT 894
 Db 221 AlaaspValThrProValAsnTyrLysTyrAlaArgThrVal 235

RESULT 10
 AAR05082
 ID AAR05082 standard; protein; 261 AA.
 XX
 AC AAR05082;
 XX
 DT 10-MAR-2003 (updated)
 DT 08-OCT-1990 (first entry)
 XX
 XX
 DE TMG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FN JP02111795-A.
 XX
 PD 24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX
 DR WPI; 1990-169109/22.
 DR N-PSDB; AAQ04687.
 XX
 XX
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX
 PS Claim 2; Fig 2; 20pp; Japanese.
 XX
 CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 SQ Sequence 261 AA;

Alignment Scores:
 Pred. No.: 7, 92e-90 Length: 261
 Score: 1142.00 Matches: 228
 Percent Similarity: 98.30% Conservatives: 3
 Best Local Similarity: 97.02% Mismatches: 4
 Query Match: 49.10% Indels: 0
 DB: 11 Gaps: 0

US-09-147-052-1 (1-1371) x AAR05082 (1-261)

QY 190 TCTATGCTATTACTTAAACGATGCAACCCCAATATGCGCAACCCCAATTAGAAGCA 249
 Db 27 CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGlnLeuGlnAla 46
 QY 250 GCGCAATGGAGTTAAGATCAATCAATGCTAAAGCGATGACATAGCTTCACTACAA 309
 Db 47 AlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaArgThrLeuAlaSerLeuGln 66

QY 310 GACTATGCCAGATTGAGCTAGCTTTATCATCTGCTTTATAGTGAAGCTGAACAGTTAAC 369
 Db 67 AspYrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsn 86
 QY 370 AATAACCTTAAATGCAACATTAGAACAACTAAATGGCTAAACTAATATTAGATCAGCC 429
 Db 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 QY 430 ATCAACCAAGCTAATACGGATTAACAGCTTTTGATAATGAACACCCAAATTTAGTTGAA 489
 Db 107 IleAsnGlnAlaAsnThrAspLysThrPheAspAsnGluHisProAsnLeuValGlu 126
 QY 490 GCATACAAAGCAGCTAAAACCCACTTTAGAACACGCTGCTACTAACCTTGAAGTTGTCA 549
 Db 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
 QY 550 TCAACTGCTTAAATCAAAATTCGCAATAATTTAGTGGATCTATACATAAAGCTAGT 609
 Db 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAsnAlaSerSer 166
 QY 610 TTAATAACTAAAACACTAGATCCACTAAATGGGGGACGCTTTTAGATTCTAATGAGATT 669
 Db 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluIle 186
 QY 670 ACTACAGCTAAATAAGAAATTAATAATACGTTTCACTTAACTATTAAATGAACAAAAGCTAAT 729
 Db 187 ThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsn 206
 QY 730 GCTGATGCAATTATCTAATAGTTTTTAAATAAGTGAATTCAAAATAATGAACAAAAGCTTT 789
 Db 207 AlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPhe 226
 QY 790 GTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAACCTTCAACAGTTTGTGCTTTTACT 849
 Db 227 ValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPheSer 246
 QY 850 GCTGATGTAACACCCGCTCAATTAATAATGAAGAGGACCGTT 894
 Db 247 AlaaspValThrProValAsnTyrLysTyrAlaArgThrVal 261

RESULT 11
 AAR79911
 ID AAR79911 standard; Protein; 261 AA.
 XX
 AC AAR79911;
 XX
 DT 19-JUL-1996 (first entry)
 XX
 DE M.gallisepticum 261 amino acid protein.
 XX
 KW Detection; probe; primer; PCR; amplification; secretion; lung;
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.
 XX
 OS Mycoplasma gallisepticum.
 XX
 PN JP07236498-A.
 XX
 PD 12-SEP-1995.
 XX
 PF 25-FEB-1994; 94JP-0052764.
 XX
 PR 25-FEB-1994; 94JP-0052764.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 DR WPI; 1995-347462/45.
 DR N-PSDB; AAT04076.
 XX
 PT Detection of Mycoplasma gallisepticum - for the quick detection,
 PT i.e. within one day, of avian chronic respiratory diseases
 XX
 PS Claim 3; Page 10-11; 11pp; Japanese.

ss;
ss;

poultry

inserted
protein

100

GAAGCA 249
:::|
GlnAla 46

CTACAA 309
|||||
LeuGln 66

ValAsn 86

370 AATAACCTTAATCAACATTAGACAACTAAATGGCTAAACACTAATTTAGATCAGCC 429
 87 AsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAla 106
 430 ATCAACCAAGCTAATACGGATAAAGACATTTGATAATGAACACCCAAATTTAGTTGAA 489
 107 IleAsnGlnAlaAsnThrAspLysThrPheAspAsnGluHisProAsnLeuValGlu 126
 490 GCATACAAACACTAAACCACTTTAGAACACGGCTGCTACTAACCTTTGAAGTTTGTCA 549
 127 AlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuAla 146
 550 TCAACTGCTTATATCAAAATTCGAATTAATTTAGTGGATCTATACATAAAGCTAGT 609
 147 SerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnAlaSerSer 166
 610 TTAATACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATT 669
 167 LeuIleThrLysThrLeuAspProLeuAsnGlyGlyMetLeuLeuAspSerAsnGluLle 186
 670 ACTACAGCTAATAAGAAATTAATATAGCTTATCACTCAACTATTAAATGAA---CAAAAGACT 726
 187 ThrThrValAsnArgAsn---AsnAsnThrLeuSerThrIleAsnGluGlnLysThr 205
 727 AATGCTGATGCATTATCTAATAGTTTATTAATAAAGTGTATCAAAATTAATGAACAACT 786
 206 AsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSer 225
 787 TTTGTAGGACGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTT 846
 226 PheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheValAlaPhe 245
 847 ACTGCTGATGTACACCCGCTCAATTAATAATATGCAAGAACGACGCTT 894
 246 SerAlaAspValThrProValAsnTyrLysTyrAlaArgThrVal 261

RESULT 13

AAR63226
 ID AAR63226 standard; Protein; 661 AA.
 AC
 AC AAR63226;
 DT 25-MAR-2003 (updated)
 DT 23-JUN-1995 (first entry)
 XX
 DE Mycoplasma gallisepticum antigen (UM-81).
 XX recombinant avipox virus; live vaccine; mycoplasma antigen.
 KW Mycoplasma gallisepticum.
 OS
 FH Key Location/Qualifiers
 FT Protein 1..661
 FT /note= "Trp residues correspond to TGA codons"
 XX
 XX W09423019-A1.
 XX
 PD 13-OCT-1994.
 XX
 XX 31-MAR-1994; 94WO-JP00541.
 XX
 PF 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX
 XX (JAP) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeiki S, Saitos, Takahashi K;
 XX
 DR WPI; 1994-333181/41.
 DR N-PSDB; AAR77853.
 XX

PT Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 of a live vaccine
 XX
 PS Claim 4; Page 61-65; 123pp; Japanese.

XX A restriction fragment of the insert of M.gallisepticum genomic clone
 CC pUM-81 containing an open reading frame was sequenced (AAQ77853). The
 CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox
 CC virus comprising the coding sequence can be used as a live vaccine to
 CC protect against infection by Mycoplasma gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 661 AA;

Alignment Scores:

Pred. No.: 6,02e-61 Length: 661
 Score: 807.00 Matches: 188
 Percent Similarity: 55.75% Conservative: 64
 Best Local Similarity: 41.59% Mismatches: 128
 Query Match: 34.69% Indels: 72
 DB: 15 Gaps: 13

US-09-147-052-1 (1-1371) x AAR63226 (1-661)

QY 151 CCACCAGTGGTTCAACCGTGCCTAGATTGGCTGTATGCTCTATTACTAAAAA 210
 Db 35 ProProSerGlyGly-----AsnMetAsnGlyGly 44
 QY 211 GATGCAACCCAAATTAATGGCCAAACCAATTAAGAGCAGCGCAATGAGTAAACAGAT 270
 Db 45 AsnThrAsnProSerAspGlyGlnGlyMetMetAsnAlaAlaLysGluLeuAlaAsp 64
 QY 271 -----CTAATCAATGCTAAAGCGATGCACATTAAGCTTCACTACAA 309
 Db 65 AlaLysAlaAlaLeuThrThrLeuIleAsnGlyGluThrAlaAsnLeuAlaSerTyrGlu 84
 QY 310 GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGCTTAAC 369
 Db 85 AspTyrAlaLysIleLysSerGluLeuThrSerAlaTyrGluThrAlaLysAlaValSer 104
 QY 370 AATAACCTTAATGCAACATTAGAACCACTAAATGGCTAAACCTAATTTAGATCAGCC 429
 Db 105 AlaLysThrGlyAlaThrLeuAsnGluValAsnGluAlaLysThrThrLeuAspAlaAla 124
 QY 430 ATCAACCAAGCTAATACGGATAAAGACGCTTTGATAATGAACACCCAAATTTAGTTGAA 549
 Db 125 IleLysLysAlaAlaSerAlaLysAsnAspPheAspAlaGlnHisGlySerLeuValGlu 144
 QY 490 GCATACAAACACTAAACCACTTTAGAACCAAGCTGCTACTAACTTGAAGTTTGTCA 549
 Db 145 AlaTyrAsnAsnLeuLysGluThrLeuLysGluGluLysThrAsnLeuAspSerLeuAla 164
 QY 550 TCAACTGCTTATATCAAAATTCGAATTAATTTAGTGGATCTATACATAAAGCTAGT 609
 Db 165 AsnGluAsnTyrAlaAlaIleArgThrAsnLeuAsnSerLeuTyrGluLysAlaAsnThr 184
 QY 610 TTAATACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATT 669
 Db 185 IleValThrAlaThrLeuAspProAlaThr---GlyAsnIleProGluValMetSerVal 203
 QY 670 ACTACAGCTAATAAGAAATTAATATAGCTTATCACTCAACTATTAAAGCAAGACTAAT 729
 Db 204 ThrGlnAlaAsnGlnAspIleThrAsnAlaThrSerArgLeuIleAlaTyrLysGlnAsn 223
 QY 730 GCTGATGCATTATCTAATAGTTTATTAATAAAGTGTATCAAAATTAATGAACAAGTTT 789
 Db 224 AlaAspAsnLeuAlaAsnSerPheIleLysGlnSerLeuValLysAsnAsnLeuThrArg 243
 QY 790 GTAGGGGACTTTTACAAACGCTAATGTTTCAACCTTCAACCTTCAACCTTGTGCTTTTGT 849
 Db 244 ValAspValAlaAsnAsnGlnGlu---GlnProAlaAsnTyrSerPheValGlyPheSer 262

```

QY 850 GCTGATGTAACACCGCTCAATATATAATATCAAGAGGACCGTTTGGAAATGGTCATGAA 909
Db 263 ValAsnValAspThrProAsnThrPasnPheAlaGlnArgLysValTrpAlaSerGluAsn 282
QY 910 -----CCTCAAGTAGAATCTTGCACAAACACGAATAGTATCATCAGAT 951
Db 283 ThrProLeuAlaThrThrProAlaGluAspAlaThrGlnGlnAlaAlaSerLeuThrAsp 302
QY 952 GTTCTTGGATTATAGTTTACGTCGGAACAACAGGAGTACCAATTTAGTTTACCAAC 1011
Db 303 ValSerTrpIleTyrSerLeuAsnGlyAlaGluAlaLysTyrThrLeuSerPheArgTyr 322
QY 1012 TATGGT---CCATCAACATGGTTATTTATATATTCCTTATAGTTGGTTAAAGCAGCTCAT 1068
Db 323 PheGlyAlaGluLysThrAlaTyrPheProTyrLysLeuValLysThrSerAsp 342
QY 1069 GCTATAACGTTGGATTACAATAACAATAATGGAATGTTCAACAAGTTGAGTTT 1128
Db 343 -----AsnValGlyLeuGlnTyrLysLeuAsnGlyGlyAspThrLysGlnIleAsnPhe 360
QY 1129 -----GCCACTCAACTAGTGCAAAT-----AATACTACAGCT 1161
Db 361 ValGlnThrProAlaSerGlySerSerAspValAlaAlaAsnGluGluThrMetAla 380
QY 1162 AATCCA-----ACTCCACGAGTTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db 381 SerProAlaGluMetGlnSerAlaProThrValAspAspIleLysIleAlaLysValAla 400
QY 1207 TTATCAGGTTTAAGATTGGCCAAACACAAATCGAATTAAGTTGCCAACGGGTGAAGGA 1266
Db 401 LeuSerAsnLeuLysPheAsnSerAsnThrIleGluPheSerValProThrGly----- 418
QY 1267 AATATGAATAAGTTGCCCAATAGATGCGCAACATTTATCTTAGTCGA----- 1314
Db 419 -----LysAlaAlaProMetIleGlyAsnMetTyrLeuThrSerSerAsnSerGlu 435
QY 1315 -----AATCAAAATAT 1326
Db 436 ValAsnLysAsnLysIleTyrAspAspLeuPheGlyAsnSerPheAsnAsnGluAsnAsn 455
QY 1327 -----GCTGATAAGATCCCGGGTAC 1347
Db 456 ProThrAlaValThrValAspLeuLeuLysGlyTyr 467

RESULT 14
AAR79910
ID AAR79910 standard; Protein; 661 AA.
XX AC AAR79910;
XX DF 19-JUL-1996 (first entry)
XX DE M.gallisepticum 661 amino acid protein.
XX KW Detection; probe; primer; PCR; amplification; secretion; lung;
XX KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX OS Mycoplasma gallisepticum.
XX FH Key
XX FT Misc-difference 220 /note= "encoded by TGA"
XX FT Misc-difference 270 /note= "encoded by TGA"
XX FT Misc-difference 305 /note= "encoded by TGA"
XX FT Misc-difference 581 /note= "encoded by TGA"
XX FT Misc-difference 581 /note= "encoded by TGA"
XX PN JP07236498-A.
XX PD 12-SEP-1995.
XX

```

```

PF 25-FEB-1994; 94JP-0052764.
XX 25-FEB-1994; 94JP-0052764.
XX (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX WPI; 1995-347462/45.
DR N-PSDB; AAT04075.
PT Detection of Mycoplasma gallisepticum - for the quick detection,
PT i.e. within one day, of avian chronic respiratory diseases
XX Claim 2; Page 7-10; 11pp; Japanese.
XX This is the amino acid sequence of a 661 amino acid protein encoded
CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
CC sequence and the sequence of AAT04076 (encoding a 261 amino acid
CC protein) can be used to detect M.gallisepticum using probes based on
CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
CC complement of bases 893-919, 1908-1934 and the complement of bases
CC 2184-2210 of the sequence, and a probe based on nucleotides 718-41 of
CC AAT04076. The method using these sequences is faster i.e. is able to
CC detect M.gallisepticum, which causes avian chronic respiratory diseases,
CC within one day, from avian secretions, washings from the lung,
CC respiratory tract, nasal cavity, etc.
XX SQ Sequence 661 AA;
Alignment Scores:
Pred. No.: 6,028-61 Length: 661
Score: 807.00 Matches: 188
Percent Similarity: 55.75% Conservative: 64
Best Local Similarity: 41.59% Mismatches: 128
Query Match: 34.69% Indels: 72
DB: 16 Gaps: 13
US-09-147-052-1 (1-1371) x AAR79910 (1-661)
QY 151 CCACGAGTGGTTCACCGCTGATCGCTAGAAATTCGGCTGATGCTATTACTAAAAA 210
Db 35 ProProSerGlyGly-----AsnMetAsnGlyGly 44
QY 211 GATGCAAAACCAATTAATGGCCAAACCAATTAGAAGCAGCGCGAATGAGTACAGAT 270
Db 45 AsnThrAsnProSerAspGlyGlnGlyMetAsnAlaAlaAlaLysGluLeuAlaAsp 64
QY 271 -----CTAATCAATGCTAAGCGGATGACATTCAGCTTCACTACAA 309
Db 65 AlaLysAlaAlaLeuThrThrLeuLeuAsnGlyGluThrAlaAsnLeuAlaSerTyrGlu 84
QY 310 GACTATGCCAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 369
Db 85 AspTyrAlaLysIleLysSerGluLeuThrSerAlaTyrGluThrAlaLysAlaValSer 104
QY 370 AATAACCTTAATGCAACATTAGAACACATAAAATGGCTAAACCTAAATTTAGATCAGCC 429
Db 105 AlaLysThrGlyAlaThrLeuAsnGluValAlaLysThrThrLeuAspAlaAla 124
QY 430 ATCAACCAAGCTAATACGGATAAAACGACTTTTGATATGAACACCAATTTAGTTGAA 489
Db 125 IleLysLysAlaAlaSerAlaLysAsnAspPheAspAlaGlnHisGlySerLeuValGlu 144
QY 490 GCATCAAAAGCATAAAACCACTTTAGAACACCGTCTACTAACCTTGAAGGTTGTGCA 549
Db 145 AlaTyrAsnAsnLeuLysGluThrLeuLysGluLysThrAsnLeuAspSerLeuAla 164
QY 550 TCAACTGCTTATATCAAAATTCGCAATATTTAGTGGATCTATACAATAAAGCTAGTAGT 609
Db 165 AsnGluAsnTyrAlaAlaIleArgThrAsnLeuAsnSerLeuTyrGluLysAlaAsnThr 184
QY 610 TTAATAACTAAACACACTAGATCCACTAAATGGGGGACCGCTTTTGTAGTTCTAATGAGATT 669

```

```

Db      185  IleValThrAlaThrLeuAspProAlaThr---GlyAsnIleProGluValMetSerVal 203
QY      670  ACTACAGCTATAGATATTAATAATACGTATATCAACTATTATGTAACAAAGAACTAAT 729
Db      204  ThrGlnAlaAsnGlnAspIleThrAsnAlaThrSerArgLeuIleAlaTrpLysGlnAsn 223
QY      730  GCTGATGCATTCATATCTATCTTTATTAATAAAGTGGATTCAAAATATATGACAAAGATTTT 789
Db      224  AlaAspAsnLeuAlaAsnSerPheIleLysGlnSerLeuValLysAsnAsnLeuThrArg 243
QY      790  GTAGGAGCTTTTACAAACGCTTAATGTTCACACCTTCAAACTACAGTTTGTGCTTTTGT 849
Db      244  ValAspValAlaAsnAsnGlnGlu---GlnProAlaAsnTyrSerPheValGlyPheSer 262
QY      850  GCTGATGTACACCCGTCATATTAATATGTCGAAGAGCCGTTTGGATGTGTGATGAA 909
Db      263  ValAsnValAspThrProAsnTrpAsnPheAlaGlnArgLysValTrpAlaSerGluAsn 282
QY      910  -----CCTTCAAGTAGAATCTTTCACAAACGATATACATATATCAGAT 951
Db      283  ThrProLeuAlaThrThrProAlaGluAspAlaThrGlnGlnAlaSerLeuThrAsp 302
QY      952  GTTCTTGGATTATAGTTAGTGGACAAACACGAGTACCAATTTAGTTTAGCAAC 1011
Db      303  ValSerTrpIleTyrSerLeuAsnGlyAlaGluAlaLysTyrThrLeuSerPheArgTyr 322
QY      1012  TATGGT---CCATCACTGCTTATATATTTTCCCTTATAGTTGGTTAAAGCAGCTGAT 1068
Db      323  PheGlyAlaGluLysThrAlaTyrLeuTyrPheProTyrLysLeuValLysThrSerAsp 342
QY      1069  GCTAATAACGTTGGATTACATAACAATAATGAATCTTCAACAGTTGAGTTT 1128
Db      343  -----AsnValGlyLeuGlnTyrLysLeuAsnGlyGlyAspThrLysGlnIleAsnPhe 360
QY      1129  -----GCCACTTCACTAGTGCAAT-----AATACACAGCT 1161
Db      361  ValGlnThrProAlaSerGlySerSerAspValAlaAlaAsnGluGluGluThrMetAla 380
QY      1162  AATCCA-----ACTCCAGCATGTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db      381  SerProAlaGluMetGlnSerAlaProThrValAspAspIleLysIleAlaLysValAla 400
QY      1207  TTATCAGGTTTAAAGATTTGGCCAAAACACAATCGAATTAAGTTTCCACGGGTGAAGGA 1266
Db      401  LeuSerAsnLeuLysPheAsnSerAsnThrIleGluPheSerValProThrGly----- 418
QY      1267  AATATCAATAAGTTCCGCCAATGATGGCACTTTATCTTAGCTCA----- 1314
Db      419  -----LysAlaAlaProMetIleGlyAsnMetTyrLeuThrSerSerAsnSerGlu 435
QY      1315  -----AATGANAATAAT 1326
Db      436  ValAsnLysAsnLysIleTyrAspAspLeuPheGlyAsnSerPheAsnAsnGluAsnAsn 455
QY      1327  -----GCTGATAGATCCCGGGTAC 1347
Db      456  ProThrAlaValThrValAspLeuLeuLysGlyTyr 467

```

RESULT 15

AAW11978

XX

AC

XX

XX

DT

XX

DE

XX

KW

KW

XX

OS

XX

Key Peptide Location/Qualifiers
1..25 /label= Sig_peptide
/note= "the signal peptide shows homology to the pmGAL.3 signal peptide"

CA2135330-A.
11-MAY-1995.
08-NOV-1994; 94CA-2135330.
20-APR-1994; 94US-0230312.
10-NOV-1993; 93AU-0050593.
(BROW/) BROWNING G F.

Browning GF, Glew MD, Markham PF, Walker ID, Whithear KG;
WPI; 1995-241027/32.
N-PSDB; AAT51531.

New promoter region from a Mycoplasma gallisepticum adhesin gene -
useful when coupled to foreign antigen gene, for prodn. of
multivalent live vaccines, also new probes for detecting Mycoplasma
and manipulating its genome

Disclosure; Fig 3; 81pp; English.

Adhesin pmGAL.2 (AAW11978) and adhesin pmGAL.3 fragment (AAW11979)
are products of gene sequences (see also AAT51531) isolated from
Mycoplasma gallisepticum. DNA constructs incorporating the
promoter and/or signal sequences of the pmGA genes can be used
in the prodn. of multivalent live vaccines. The signal peptide
sequence is utilised where attachment of an exogenous antigen
gene to the mycoplasma cell membrane is required.

Sequence 647 AA;

Alignment Scores:

Pred. No.: 4,39e-53 Length: 647
Score: 715.50 Matches: 174
Percent Similarity: 53.38% Conservative: 63
Best Local Similarity: 39.19% Mismatches: 134
Query Match: 30.76% Indels: 73
DB: 16 Gaps: 13

US-09-147-052-1 (1-1371) x AAW11978 (1-647)

QY 157 GTGGGTTCAACGTCGTCGTAGATTTCGGCTGTATCTCTATTACTATAAAAA----- 210
Db 15 IleGlySerPheValMetLeuAlaAlaLaserCysThrThrProThrProAsnProThr 34
QY 211 -----GATGCAACCAATAATGCG 231
Db 35 ProAsnProAsnProProSerGlyGlyMetAsnGlyGlyAspThrAsnProGlyAspGly 54
QY 232 CAA-----ACCAATTAGAAGCAGCGCGAATGAGTGAACAGAT 270
Db 55 GlnGlyMetMetAsnAlaAlaSerGlnGluLeuAlaAlaAlaArgMetGlyLeuThrThr 74

QY 271 CTAATCAATGCTAAAGCGATGACATTAGCTTCACATACAGACTATGCCAAGATTGAAGCT 330
Db 75 IlePheAspSerLysAlaLysAsnLeuGlyLeuTyrValAspTyrLysLysThrGlnAsn 94
QY 331 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAACAATTAACCTTAATGCAACATTA 390
Db 95 ThrLeuThrLysAlaTyrAspAlaLysThrValLeuAspAsnSerSerSerThrThr 114

QY 391 GAACAACATAAAATGGCTAAACTAATTTAGTAATCAGCCATCAACCAAGCTAATACGAT 450
Db 115 GlnLysLeuAsnGluAlaLysThrArgLeuGluThrAlaIleArgThrAlaLysThrSer 134


```
QY 451 AAAACGACCTTTGATTAATCAACACCCAAATTTAGTTGAAGCATACAAACGACATAAAACC 510
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 LysGlnThrPheAspGluGlnHisAlaGluLeuValLysValTyrLysGluLeuLysThr 154
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
511 ACTTTAGAACAAACGGTCTACTACCTTGAAGGTTGTGTCATCACTCAACTGCTTATATCAAT 570
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 ThrLeuSerAsnGluThrAlaThrLeuAlaProTyrAlaAspAlaGlnTyrAlaGlyIle 174
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
571 CGCAATAATTTAGTGGATCTATACATAAAGCTAGTAGTTTAACTAAACACACTAGAT 630
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 LysMetHisLeuSerGlyLeuTyrAspAlaGlyLysAlaIleThrThrLysThrLeuGlu 194
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 CCACATAAATGGGGACGGCTTTAGATTCTAATGAGATTACTACACCTAATAAGATATT 690
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 ProValGluGlyAspPro---LeuThrAlaSerAlaValMetMetAlaAsnThrLysIle 213
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 AATAATACGTTA-----TCAACTATTATGAACAAAGACTAAT-----GCT 732
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 ValGluAlaIleLysAspGluValLeuAsnProGlnLysGluAsnAlaThrLysLeuAla 233
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
733 GATGATTATCTAATAGTTTATTAATAAAGTGAATCAAAATAAATGAACAAAGTTTGTGA 792
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 AspSerLeuLeuSerSerIleValLysLysIleThrGlyValGluGlu----- 249
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
793 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTCTTTTACTGCT 852
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 -----AlaHisAsnLysAlaGlnProAlaAsnTyrSerPheValGlyTyrLysArg 266
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
853 GATGTAACACCGTC-----AATTATAAATATGCAAGAGG 888
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 TrpTyrThrGluLeuLeuLeuAspLysGlnValPheProAsnTrpAspTyrAlaGluArg 286
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
889 ACCGTTTGG---AATGGTGATGAACCTTCAAGTAGAATTTGCAACACAG----- 936
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
287 ThrIlePheThrAsnSerAspGluPro-----ArgSerIleSerAsnThrProAlaAsp 304
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
937 -----AATAGTATCACAGATGTTCTTGGATTATAGTTTACTGGGAACA 981
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 GlyGlnThrMetAlaGlnProLeuSerAsnValSerTrpIleTyrSerLeuAlaGlyThr 324
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
982 AACACGAAGTACCAATTTAGTTTACCAACTATGTCCTCACTCACTGGTTATTATATTC 1041
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 GlyAlaLysTyrThrLeuGluPheThrTyrTyrGlyProSerThrGlyTyrLeuTyrPhe 344
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1042 CCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGATTGATTAACATAAATAAT 1101
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
345 ProTyrLysLeuValAsnThrSerAspGlnValLysLeuGlyLeuGluTyrLysLeuAsn 364
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1102 AATGGAATGTTCAACAAGTTGCTGCTTGCCTTCACTTCACTAGTCCA----- 1146
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 Asp-----AlaThrLysProSerAlaIleThrPheGlySer 376
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1147 AATAATACTACAGTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTT 1206
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
377 AspGlnThrMetAsnGlyLysThrProThrValAsnAspIleAsnValAlaLysValThr 396
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1207 TTATCAGGTTTAAGATTTCGCCAAACACAATCGAATTAAAGTGTTCACACGGGTGAAGGA 1266
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
397 LeuAlaAsnLeuAsnPheGlySerAsnLysIleGluPheSerValProAla----- 413
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1267 AATATGAATAAAGTTCGCCCAATGATTGGCAACATTTATCTTAGCTCAATAAATAAT 1326
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 -----GluLysValSerProMetIleGlyAsnMetTyrLeuSerSerProAsnAsn 431
QY      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1327 GCTGATAAGATC 1338
Db      ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
432 TrpAsnLysIle 435
```

Search completed: October 8, 2003, 17:52:08
Job time : 93.1459 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:32:42 ; Search time 3.545 Seconds
(without alignments)
5641.626 Million cell updates/sec

Title: US-09-147-052-1_COPY_1_63

Perfect score: 115

Sequence: 1 atgcactattttagcgga.....ttctatatgtgtaacgaatca 63

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xl
-O=/cgn2_1/USPTO_spool_US09147052/runat_08102003_154340_29792/app_query.fasta_1.5980
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0 -1-LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cd1
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052_ECGN_1_1_474_Erunat_08102003_154340_29792 -NCPU=6 -ICPU=3
-NO_MMMap -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	115	100.0	456	18	AAW36050	Hybrid Marek's dis
2	115	100.0	865	14	AAR30169	Marek's disease VI
3	115	100.0	1086	18	AAW36051	Hybrid Marek's dis
4	58	50.4	61	23	ABP39693	Staphylococcus epi
5	53	46.1	582	23	ABP73516	Candida albicans e
6	48	41.7	276	23	ABP41996	Human ovarian anti
7	48	41.7	317	20	AAW81968	Mouse E124 protein
8	48	41.7	318	20	AAW81969	Mouse E124 protein
9	48	41.7	340	19	AAW61615	Human apoptosis-as
10	48	41.7	342	19	AAW75049	Fragment of human
11	48	41.7	342	23	ABG95555	Human novel secret
12	47	40.9	69	23	ABB81404	Human prostate spe
13	47	45.2	217	19	AAW69656	Human eukaryotic-1
14	47	45.2	217	19	AAW69655	Human eukaryotic-1
15	47	45.2	217	19	AAW69657	Human eukaryotic-1
16	47	45.2	217	19	AAW69658	Human eukaryotic-1
17	47	45.2	217	20	AAW94251	Murine eukaryotic
18	47	45.2	217	21	AAW78503	Murine general tra
19	47	45.2	217	21	AAW78505	Human general tran
20	47	45.2	281	24	ABU11755	Human MDDT polyep
21	47	40.9	2958	22	ABB71368	Drosophila melanog
22	46	40.0	1641	20	AAW05845	Banana ripening fr
23	45	39.1	43	24	ABR01006	Human gene 60-enco
24	45	39.1	43	24	ABP95529	Human secreted pro
25	45	39.1	44	21	AAW27794	Human secreted pro
26	45	39.1	110	19	AAW77631	Staphylococcus aur
27	45	39.1	135	20	AAW37865	Amino acid sequenc
28	45	39.1	456	23	ABP38814	Staphylococcus epi
29	45	43.3	463	24	ABP99428	Amino acid sequenc
30	45	39.1	968	21	RAY78946	Polycystic kidney
31	45	39.1	968	22	ABG68450	Amino acid sequenc
32	45	43.3	1218	22	ABB63445	Drosophila melanog
33	44.5	38.7	63	21	AAW30061	Arabidopsis thalia
34	44.5	38.7	866	22	AAW68448	Amino acid sequenc
35	44	38.3	202	23	AAE18454	Fibroblast succin
36	44	38.3	216	18	AAW18232	Fibroblast succin
37	44	42.3	265	24	ABU00869	S. pneumoniae type
38	44	38.3	269	18	AAW18233	Fibroblast succin
39	44	38.3	608	18	AAW18230	Fibroblast succin
40	44	38.3	608	21	AAW48551	Fibroblast succin
41	44	38.3	608	23	AAO18657	F succinogenes xyl
42	44	38.3	1550	24	ABG72697	Human ATP-binding
43	44	38.3	1873	22	AAU04484	Human PD-ATP-bind
44	44	38.3	2059	24	ABU54629	Human NOVX polyep
45	44	38.3	2144	23	AAU09174	Human transporter

ALIGNMENTS

RESULT 1

AAW36050

ID AAW36050 standard; Protein; 456 AA.

XX

AC AAW36050;

XX

DT 15-JUL-1998 (first entry)

XX

DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX

KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

KW antigen; vaccine; poultry.

XX

OS Chimeric - Marek's disease gammaherpesvirus.

OS Chimeric - Mycoplasma gallisepticum.

XX

FT Key Location/Qualifiers

Region 1..64

FT /note= "derived from Marek's disease virus gB protein"
 FT 65..456
 FT /note= "derived from M. gallisepticum antigenic protein"
 PN WO9736924-A1.
 XX 09-OCT-1997.
 XX 28-MAR-1997; 97WO-JP01084.
 XX 29-MAR-1996; 96JP-0103548.
 XX (JAPG) NIPPON ZEON KK.
 PA Saito S, Tsuzaki Y, Yanagida N;
 PI WPI; 1997-503046/46.
 XX N-PSDB; AAT96595.
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX Disclosure; Page 16-19; 51pp; Japanese.
 XX This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX SQ Sequence 456 AA;

Alignment Scores:
 Pred. No.: 7.89e-11 Length: 456
 Score: 115.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 18 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x AAW36050 (1-456)

QY 1 ATGCATATTTAGCGGAATTCATATTTTCCTTATAGTATTCATATGTCAGAAC 60
 Db 1 MethIstYrPheArgAsnCysIlePhePheLeuValIleLeuTyrglyThrAsn 20
 QY 61 TCA 63
 Db 21 Ser 21

RESULT 2
 AAR30169
 ID AAR30169 standard; Protein; 865 AA.
 XX AAR30169;
 XX 25-MAR-2003 (updated)
 DT 07-MAY-1993 (first entry)
 XX Marek's Disease Virus glycoprotein B homologue of HSV.

DE Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
 KW Herpes Simplex Virus.
 XX Marek's Disease Virus.
 OS EP520753-A1.
 PN 30-DEC-1992.
 PD 24-JUN-1992; 92EP-0305775.

XX 28-JUN-1991; 91US-0722860.
 PR 10-DEC-1991; 91US-0803633.
 XX (JAPG) NIPPON ZEON KK.
 PA (USDA) US SEC OF AGRIC.
 XX Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
 PI WPI; 1993-001546/01.
 XX Recombinant fowl pox virus contg. Marek's disease virus antigen
 PT gene - used to produce cell-free vaccine against Marek's disease
 PT virus
 XX Example 2; Page 15-19; 30pp; English.
 XX The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
 CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-SalI
 CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
 CC from K3 fragment were ligated with EcoRI, SalI digested pUC18. The
 CC sequence of the putative MDV gBh was determined by sequencing a set
 CC of deletion mutants. The nucleotide and amino acid sequences were
 CC found to be identical with the published sequences of the gBh of
 CC RBIB strain of MDV (Ross et al., J.Gen.Virol., 70:1789-1894, 1988).
 CC A fragment contg. the entire coding region of MDV gBh was inserted
 CC into pNZ1729R (see AAQ34774-Q34778) to produce a recombinant FPV/MDVgBh
 CC virus for immunising chickens. See also AAQ34780-Q34781.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 865 AA;

Alignment Scores:

Pred. No.: 8.38e-11 Length: 865
 Score: 115.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x AAR30169 (1-865)

QY 1 ATGCATATTTAGCGGAATTCATATTTTCCTTATAGTATTCATATGTCAGAAC 60
 Db 1 MethIstYrPheArgAsnCysIlePhePheLeuValIleLeuTyrglyThrAsn 20
 QY 61 TCA 63
 Db 21 Ser 21

RESULT 3
 AAW36051
 ID AAW36051 standard; Protein; 1086 AA.
 XX AAW36051;
 XX 15-JUL-1998 (first entry)
 DT Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 DE Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 XX Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX Key Location/Qualifiers
 FH Region 1..672
 FT /note= "derived from Marek's disease virus gB protein"
 FT 693..1086
 FT /note= "derived from M. gallisepticum antigen"
 XX WO9736924-A1.

Constructing strains for identifying gene products as effective targets for therapeutic intervention, by inactivating in the strain one allele

DE Mouse EI24 protein.
 XX
 KW EI24; etoposide-induced apoptosis; murine; degenerative disorder;
 KW cell proliferation; cell death; p53.
 XX
 OS Mus sp.
 XX
 PN US5843659-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 21-MAR-1996; 96US-0619362.
 XX
 PR 21-MAR-1996; 96US-0619362.
 XX
 XX (APOP-) APOPTOSIS TECHNOLOGY INC.
 XX
 PI Guild BC, Lehar SM;
 XX
 DR WPI; 1999-044568/04.
 DR N-PSDB; AAV65036.
 XX
 XX Cloned etoposide-induced apoptosis gene EI24 - useful for diagnosing
 PT degenerative disorders characterised by inappropriate cell
 PT proliferation or death
 XX
 PS Claim 2; Fig 2; 35pp; English.
 XX
 CC This sequence represents a murine etoposide-induced apoptosis gene, EI24.
 CC The EI24 gene product can be used in the diagnosis of degenerative
 CC disorders characterised by inappropriate cell proliferation or death.
 CC Induction of this gene by etoposide requires expression of wild-type
 CC p53.
 XX
 SQ Sequence 317 AA;
 XX
 Alignment Scores:
 Pred. No.: 18.2 Length: 317
 Score: 48.00 Matches: 10
 Percent Similarity: 66.67% Conservative: 2
 Best Local Similarity: 55.56% Mismatches: 6
 Query Match: 41.74% Indels: 0
 DB: 20 Gaps: 0
 XX
 US-09-147-052-1_COPY_1_63 (1-63) x AAW81968 (1-317)
 QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTA 55
 DB 93 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTrpArgVal 110
 XX
 RESULT 8
 AAW81969
 ID AAW81969 standard; Protein; 318 AA.
 XX
 AC AAW81969;
 XX
 DT 05-FEB-1999 (first entry)
 XX
 DE Mouse EI24 protein #2.
 XX
 KW EI24; etoposide-induced apoptosis; murine; degenerative disorder;
 KW cell proliferation; cell death; p53.
 XX
 OS Mus sp.
 XX
 PN US5843659-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 21-MAR-1996; 96US-0619362.
 XX
 PR 21-MAR-1996; 96US-0619362.
 XX
 XX

PA (APOP-) APOPTOSIS TECHNOLOGY INC.
 XX
 PI Guild BC, Lehar SM;
 XX
 DR WPI; 1999-044568/04.
 XX
 XX Cloned etoposide-induced apoptosis gene EI24 - useful for diagnosing
 PT degenerative disorders characterised by inappropriate cell
 PT proliferation or death
 XX
 PS Example; Fig 3; 35pp; English.
 XX
 CC This sequence represents a murine etoposide-induced apoptosis gene, EI24.
 CC The EI24 gene product can be used in the diagnosis of degenerative
 CC disorders characterised by inappropriate cell proliferation or death.
 CC Induction of this gene by etoposide requires expression of wild-type
 CC p53.
 XX
 SQ Sequence 318 AA;
 XX
 Alignment Scores:
 Pred. No.: 18.2 Length: 318
 Score: 48.00 Matches: 10
 Percent Similarity: 66.67% Conservative: 2
 Best Local Similarity: 55.56% Mismatches: 6
 Query Match: 41.74% Indels: 0
 DB: 20 Gaps: 0
 XX
 US-09-147-052-1_COPY_1_63 (1-63) x AAW81969 (1-318)
 QY 2 TGCACATATTTAGCGGAATTCATATTTTCCTTATAGTTATCTATATGTA 55
 DB 94 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTrpArgVal 111
 XX
 RESULT 9
 AAW61615
 ID AAW61615 standard; Protein; 340 AA.
 XX
 AC AAW61615;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Human apoptosis-associated protein.
 XX
 KW Human; apoptosis-associated protein; cancer; autoimmune disorder;
 KW viral infection; neurodegenerative disorder; myelodysplastic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO9832854-A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 26-JAN-1998; 98WO-US01421.
 XX
 PR 29-JAN-1997; 97US-0790572.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Goli SK, Hillman JL;
 XX
 DR WPI; 1998-427950/36.
 DR N-PSDB; AAV45340.
 XX
 XX Human apoptosis-associated protein, NHAAP - useful for, e.g.
 PT treating diseases involving increased or decreased apoptosis such as
 PT cancer and auto-immune disorder(s)
 XX
 PS Claim 1; Fig 1; 59pp; English.
 XX
 CC The human apoptosis-associated protein (NHAAP), it's agonists and
 CC nucleotides encoding it can be used to treat or prevent disorders
 CC associated with decreased apoptosis, e.g. cancer (including lymphomas and

CC leukaemia) , autoimmune disorders (including systemic lupus
CC erythematosus, scleroderma and arthritis) and viral infections (including
CC HIV, herpes virus and adenovirus infection). Antisense constructs, the
CC antibodies and antagonists of NHAP can be used to treat disorders
CC associated with increased apoptosis, e.g. neurodegenerative disorders
CC (including Alzheimer's disease, Parkinson's disease and amyotrophic
CC lateral sclerosis), myelodysplastic disorders (including ischaemic
CC injury, aplastic anaemia, trauma, heart attack and AIDS).
XX
SQ Sequence 340 AA;

Alignment Scores:
Pred. No.: 18.3 Length: 340
Score: 48.00 Matches: 10
Percent Similarity: 65.67% Conservative: 2
Best Local Similarity: 55.56% Mismatches: 6
Query Match: 41.74% Indels: 0
DB: 19 Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x AAW61615 (1-340)

QY 2 TGCATATTTAGCGGAAATGCAATATTTTCCTTATAGTTATCTATATGGTA 55
Db 75 CysAlaTrpAsnGlyGlyValPheTrpPheSerLeuLeuPheTyrArgVal 92
||| |||||::: ::||| ||||| ||||| |||
||| |||||::: ::||| ||||| ||||| |||

RESULT 10

AAW75049

ID AAW75049 standard; Protein; 342 AA.

AC AAW75049;

XX

XX

DT 25-JAN-1999 (first entry)

XX

DE Fragment of human secreted protein encoded by gene 114.

XX

KW Human; secreted protein; testis; tumour; foetal brain tissue;

KW fusion protein; cancer; central nervous system; seizure;

KW diagnosis; neurodegenerative disease.

XX

OS Homo sapiens.

XX

PN WO9839448-A2.

XX

PD 11-SEP-1998.

XX

PF 06-MAR-1998; 98WO-US04493.

XX

PR 02-OCT-1997; 97US-0061060.

PR 07-MAR-1997; 97US-0038621.

PR 07-MAR-1997; 97US-0040161.

PR 07-MAR-1997; 97US-0040162.

PR 07-MAR-1997; 97US-0040163.

PR 07-MAR-1997; 97US-0040333.

PR 07-MAR-1997; 97US-0040334.

PR 07-MAR-1997; 97US-0040336.

PR 07-MAR-1997; 97US-0040336.

PR 11-APR-1997; 97US-0043311.

PR 11-APR-1997; 97US-0043312.

PR 11-APR-1997; 97US-0043313.

PR 11-APR-1997; 97US-0043314.

PR 11-APR-1997; 97US-0043358.

PR 11-APR-1997; 97US-0043359.

PR 11-APR-1997; 97US-0043576.

PR 11-APR-1997; 97US-0043578.

PR 11-APR-1997; 97US-0043580.

PR 11-APR-1997; 97US-0043669.

PR 11-APR-1997; 97US-0043670.

PR 11-APR-1997; 97US-0043671.

PR 11-APR-1997; 97US-0043672.

PR 23-MAY-1997; 97US-0043674.

PR 23-MAY-1997; 97US-0047492.

PR 23-MAY-1997; 97US-0047500.

PR 23-MAY-1997; 97US-0047501.

PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.

XX	07-MAR-1997;	97US-038621P;
PR	07-MAR-1997;	97US-040161P;
PR	07-MAR-1997;	97US-040162P;
PR	07-MAR-1997;	97US-040163P;
PR	07-MAR-1997;	97US-040333P;
PR	07-MAR-1997;	97US-040334P;
PR	07-MAR-1997;	97US-040336P;
PR	07-MAR-1997;	97US-040626P;
PR	11-APR-1997;	97US-043111P;
PR	11-APR-1997;	97US-043312P;
PR	11-APR-1997;	97US-043313P;
PR	11-APR-1997;	97US-043314P;
PR	11-APR-1997;	97US-043315P;
PR	11-APR-1997;	97US-043568P;
PR	11-APR-1997;	97US-043569P;
PR	11-APR-1997;	97US-043576P;
PR	11-APR-1997;	97US-043578P;
PR	11-APR-1997;	97US-043580P;
PR	11-APR-1997;	97US-043699P;
PR	11-APR-1997;	97US-043670P;
PR	11-APR-1997;	97US-043671P;
PR	11-APR-1997;	97US-043672P;
PR	11-APR-1997;	97US-043674P;
PR	23-MAY-1997;	97US-047492P;
PR	23-MAY-1997;	97US-047500P;
PR	23-MAY-1997;	97US-047501P;
PR	23-MAY-1997;	97US-047502P;
PR	23-MAY-1997;	97US-047503P;
PR	23-MAY-1997;	97US-047581P;
PR	23-MAY-1997;	97US-047582P;
PR	23-MAY-1997;	97US-047583P;
PR	23-MAY-1997;	97US-047584P;
PR	23-MAY-1997;	97US-047585P;
PR	23-MAY-1997;	97US-047586P;
PR	23-MAY-1997;	97US-047587P;
PR	23-MAY-1997;	97US-047588P;
PR	23-MAY-1997;	97US-047589P;
PR	23-MAY-1997;	97US-047590P;
PR	23-MAY-1997;	97US-047592P;
PR	23-MAY-1997;	97US-047593P;
PR	23-MAY-1997;	97US-047594P;
PR	23-MAY-1997;	97US-047595P;
PR	23-MAY-1997;	97US-047596P;
PR	23-MAY-1997;	97US-047597P;
PR	23-MAY-1997;	97US-047598P;
PR	23-MAY-1997;	97US-047599P;
PR	23-MAY-1997;	97US-047600P;
PR	23-MAY-1997;	97US-047601P;
PR	23-MAY-1997;	97US-047612P;
PR	23-MAY-1997;	97US-047613P;
PR	23-MAY-1997;	97US-047614P;
PR	23-MAY-1997;	97US-047615P;
PR	23-MAY-1997;	97US-047617P;
PR	23-MAY-1997;	97US-047618P;
PR	23-MAY-1997;	97US-047632P;
PR	23-MAY-1997;	97US-047633P;
PR	06-JUN-1997;	97US-048964P;
PR	06-JUN-1997;	97US-048974P;
PR	08-JUL-1997;	97US-049610P;
PR	08-JUL-1997;	97US-051926P;
PR	16-JUL-1997;	97US-052874P;
PR	18-AUG-1997;	97US-055724P;
PR	22-AUG-1997;	97US-056630P;
PR	22-AUG-1997;	97US-056631P;
PR	22-AUG-1997;	97US-056632P;
PR	22-AUG-1997;	97US-056633P;
PR	22-AUG-1997;	97US-056637P;
PR	22-AUG-1997;	97US-056622P;
PR	22-AUG-1997;	97US-056644P;
PR	22-AUG-1997;	97US-056845P;
PR	22-AUG-1997;	97US-056862P;
PR	22-AUG-1997;	97US-056864P;

US-09-147-052-1_COPY_1_63 (1-63) x AAW75049 (1-342)

RESULT 11

15-JAN-2003 (first entry)

Human; secreted protein; autoimmune disease; chemotaxis;
 KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;
 KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;
 KW nervous system disorders; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; tissue regeneration;
 KW epithelial cell proliferation; organ transplantation; food additive;
 KW preservative; nutritional.

XX
XXXX
PD
16-III-2002

PF 08-SEP-1998; 98US-0149476.

PR 22-AUG-1997; 97US-056872P.
 PR 22-AUG-1997; 97US-056874P.
 PR 22-AUG-1997; 97US-056875P.
 PR 22-AUG-1997; 97US-056876P.
 PR 22-AUG-1997; 97US-056877P.
 PR 22-AUG-1997; 97US-056878P.
 PR 22-AUG-1997; 97US-056879P.
 PR 22-AUG-1997; 97US-056880P.
 PR 22-AUG-1997; 97US-056881P.
 PR 22-AUG-1997; 97US-056882P.
 PR 22-AUG-1997; 97US-056884P.
 PR 22-AUG-1997; 97US-056886P.
 PR 22-AUG-1997; 97US-056887P.
 PR 22-AUG-1997; 97US-056888P.
 PR 22-AUG-1997; 97US-056889P.
 PR 22-AUG-1997; 97US-056892P.
 PR 22-AUG-1997; 97US-056893P.
 PR 22-AUG-1997; 97US-056903P.
 PR 22-AUG-1997; 97US-056908P.
 PR 22-AUG-1997; 97US-056909P.
 PR 22-AUG-1997; 97US-056910P.
 PR 22-AUG-1997; 97US-056911P.
 PR 22-AUG-1997; 97US-057761P.
 PR 05-SEP-1997; 97US-057650P.
 PR 05-SEP-1997; 97US-057669P.
 PR 12-SEP-1997; 97US-058785P.
 PR 02-OCT-1997; 97US-061060P.
 PR 06-MAR-1998; 98WO-US04493.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC, Bednarik DR;
 PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrrie AM;
 PI Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
 XX WPI; 2002-634796/68.
 XX
 PT New isolated human secreted protein for diagnosing, preventing,
 PT treating or ameliorating medical conditions and used as a food additive
 PT or preservative -
 XX
 PS Disclosure; Column 98-99; 129pp; English.
 XX
 CC The invention relates to an isolated protein that is one of 186 human
 CC secreted proteins, given in the specification, encoded by one of
 CC 309 cDNA sequences also given in the specification. The protein is used
 CC in a pharmaceutical composition used to prevent, treat or ameliorate a
 CC medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional
 CC components. The present sequence represents one of the novel human
 CC secreted proteins of the invention.
 XX
 SQ Sequence 342 AA;
 Alignment Scores: 342 AA;
 Pred. No.: 18.3 Length: 342
 Score: 48.00 Matches: 10
 Percent Similarity: 66.67% Conservative: 2
 Best Local Similarity: 55.56% Mismatches: 6

Query Match: 41.74% Indels: 0
 DB: 23 Gaps: 0
 US-09-147-052-1_COPY_1_63 (1-63) x ABG95555 (1-342)
 QY 2 TGCACATATTTAGCGGAATTCACATATTTTCCCTTATAGTTATCTATATGGTA 55
 ||| |||||::: ::||| ||||| ||||| |||
 Db 77 CysAlaTrpAsnGlyValPheIrpPheSerLeuLeuPheTyrArgVal 94
 RESULT 12
 ABB81404
 ID ABB81404 standard; Protein; 69 AA.
 XX
 AC ABB81404;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE Human prostate specific protein (PSP) SEQ ID NO:176.
 XX
 KW Human; prostate specific nucleic acid; prostate specific protein; PSP;
 KW PSNA; prostate cancer; prostate related disorder; prostate disease;
 KW cytostatic; gene therapy; antisense therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200242499-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 20-NOV-2001; 2001WO-US45179.
 XX
 PR 21-NOV-2000; 2000US-252186P.
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Salceda S, Macina RA, Recipon H, Cafferkey R, Ali S, Sun Y;
 PI Liu C;
 XX
 DR WPI; 2002-500285/53.
 XX
 PT New prostate specific polypeptide for detecting, diagnosing,
 PT monitoring, treating, staging and predicting cancers in humans having
 PT cancer and non-cancerous prostate disease -
 XX
 PS Claim 11; Page 239; 255pp; English.
 CC
 CC ABB89000 to ABB89111 represent human prostate specific nucleic acid
 CC (PSNA) sequences which encode the human prostate specific proteins (PSP)
 CC given in ABB81341 to ABB81439. PSNA and PSP sequences have cytostatic
 CC activity, and can be used in gene therapy, antisense therapy and
 CC vaccines. PSNA and PSP sequences can be used for diagnosing and
 CC monitoring the presence and metastases of prostate cancer in a patient.
 CC PSNA and PSP sequences are also useful: (i) for determining non-cancerous
 CC prostate disease, by measuring their expression levels and/or structural
 CC alterations; (ii) for determining that the sample that has prostate
 CC tissue-like characteristics or is a prostate tissue; (iii) as elements
 CC in an array or a multi-analyte test to recognise expression patterns
 CC associated with prostate cancer and other prostate related disorders;
 CC and (iv) as elements in a computer program for pattern recognition of
 CC prostate disorders.
 XX
 SQ Sequence 69 AA;
 Alignment Scores:
 Pred. No.: 23.4 Length: 69
 Score: 47.00 Matches: 7
 Percent Similarity: 63.16% Conservative: 5
 Best Local Similarity: 36.84% Mismatches: 7
 Query Match: 40.87% Indels: 0
 DB: 23 Gaps: 0
 US-09-147-052-1_COPY_1_63 (1-63) x ABB81404 (1-69)

Qy 39 TATAAGGAAAAATATGCAATTTCGCCCTAAATAGTG 4
|||||
Dd 183 TvrLvsGluArgLeuGluVleuProPolsIleVal 194
|||||

DO
RESULT 14

RESULT 14
AAW69655
ID AAW69655 standard: Protein: 217 AA:

ID AAW69655
XX
AC AAW69655:

AC AAW69035;
XX
DT 19-OCT-1998 (first entry)

DT 19-OCT-1998 (first entry)
 XX
 DE Human eukaryotic-initiation factor 4E variant N118A.

XX Human eukaryotic-initiation factor 4E variant N118A.
KW Human; eukaryotic-initiation factor 4E; eIF-4E; variant; mutant;
KW m7G-RNA; breast cancer; temporal regulation; differentiation;
KW stem cell; immunisation.

XX	stem cell; I
OS	Homo sapiens

OS Homo sapiens.
OS Synthetic.

XX PN WO9828325-A1

XX	2	1000
XX	2	1000

XX
PF 19-DEC-1997; 97WO-US23428

FF 19-DEC-1997; 9/WO-US233428;
XX
PR 20-DEC-1996; 96US-00333533.

PR 20-DEC-1996; 96US
XX
PA (PTVEM-) PNTV EMORY

PA (UYEM-) UNIV EMORY.
XX

PI Hagedorn CH, Spivak
XX

DR WPI; 1998-377592/32.

	XX
Purified eukaryotic induction factor 4E and its variants with	PT

PT Purified eukaryotic induction factor 4E and its variants with
PT altered affinity for capped RNA - used to alter expression levels in
PT cells, e.g. for treating breast cancer, and for developing efficient
PT transient expression system based on RNA

PS Claim 4: Page -- 31pp: English

PS The present sequence represents a specifically claimed variant human
XX eukaryotic-initiation factor 4E (eIF-4E) protein. The sequence was
CC derived from wild-type human eIF-4E as stated in the specification.
CC The variant has an altered binding affinity for m7G-RNA. DNA encoding
CC the variant eIF-4E can be used to alter expression of genes in cells,
CC e.g. for treating breast cancer where it encodes a variant of reduced
CC affinity and this acts as a dominant negative mutant, counteracting
CC overexpression of wild-type eIF-4E that maintains the cancer phenotype,
CC also for temporal regulation for controlling differentiation of stem
CC cells or expression of antigens for immunisation. A complex of eIF-4E
CC or variant eIF-4E and m7G-RNA is used to transfect cells for transient
CC expression of the RNA component. These processes are carried out in
CC vitro or in vivo. Also antibodies raised against eIF-4E or variant
CC eIF-4E are used to detect or purify the corresponding antigens.
CC Transfection with RNA leads to transient expression of a protein without
CC causing genetic alteration of the host cell. By forming a complex with
CC capped RNA, eIF-4E or variant eIF-4E improves its stability (and thus
CC transfection efficiency), resulting in immediate and efficient
CC translation to protein.

CC Translation to pro
XX
SQ Sequence 217 AA:

SV	Sequence	217 AM	
Alignment Scores:			
	Pred. No.:	26	Length: 217
	Score:	47.00	Matches: 8
	Percent Similarity:	83.33%	Conservative: 2
	Best Local Similarity:	66.67%	Mismatches: 2
	Query Match:	45.19%	Indels: 0
	DB:	19	Gaps: 0

US-09-147-052-1_COPY_1_63 (1-63) x AAW69655 (1-217)

QY 39 TATAAGGAAATATGCAATTCGCTAAATAGTG 4
 Db 183 TyrLysGluArgLeuGlyLeuProProlysIleVal 194

RESULT 15

AAW69657

ID AAW69657 standard; Protein; 217 AA.

XX AC AAW69657;

XX 19-OCT-1998 (first entry)

XX Human eukaryotic-initiation factor 4E variant I115A.

XX Human; eukaryotic-initiation factor 4E; eIF-4E; variant; mutant;
 KW m7G-RNA; breast cancer; temporal regulation; differentiation;
 KW stem cell; immunisation.

XX Homo sapiens.

XX Synthetic.

XX PN W09828325-A1.

XX PD 02-JUL-1998.

XX PF 19-DEC-1997; 97WO-US23428.

XX PR 20-DEC-1996; 96US-0033533.

XX PA (UYEM-) UNIV EMORY.

XX PI Hagedorn CH, Spivak-Kroizman T, Xie Y;

XX DR WPI; 1998-377592/32.

XX Purified eukaryotic induction factor 4E and its variants with
 PT altered affinity for capped RNA - used to alter expression levels in
 PT cells, e.g. for treating breast cancer, and for developing efficient
 PT transient expression system based on RNA

XX Claim 7; Page -: 31pp; English.

XX The present sequence represents a specifically claimed variant human
 CC eukaryotic-initiation factor 4E (eIF-4E) protein. The sequence was
 CC derived from wild-type human eIF-4E as stated in the specification.
 CC The variant has an altered binding affinity for m7G-RNA. DNA encoding
 CC the variant eIF-4E can be used to alter expression of genes in cells,
 CC e.g. for treating breast cancer where it encodes a variant of reduced
 CC affinity and this acts as a dominant negative mutant, counteracting
 CC overexpression of wild-type eIF-4E that maintains the cancer phenotype,
 CC also for temporal regulation for controlling differentiation of stem
 CC cells or expression of antigens for immunisation. A complex of eIF-4E
 CC or variant eIF-4E and m7G-RNA is used to transfect cells for transient
 CC expression of the RNA component. These processes are carried out in
 CC vitro or in vivo. Also antibodies raised against eIF-4E or variant
 CC eIF-4E are used to detect or purify the corresponding antigens.
 CC Transfection with RNA leads to transient expression of a protein without
 CC causing genetic alteration of the host cell. By forming a complex with
 CC capped RNA, eIF-4E or variant eIF-4E improves its stability (and thus
 CC transfection efficiency), resulting in immediate and efficient
 CC translation to protein.

XX SQ Sequence 217 AA;

Alignment Scores:

Pred. No.:	26	Length:	217
Score:	47.00	Matches:	8
Percent Similarity:	83.33%	Conservative:	2
Best Local Similarity:	66.67%	Mismatches:	2
Query Match:	45.19%	Indels:	0
DB:	19	Gaps:	0

US-09-147-052-1_COPY_1_63 (1-63) x AAW69657 (1-217)
 QY 39 TATAAGGAAATATGCAATTCGCTAAATAGTG 4
 Db 183 TyrLysGluArgLeuGlyLeuProProlysIleVal 194

Search completed: October 8, 2003, 17:52:10
 Job time : 5.545 secs

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:50:40 ; Search time 48.58 Seconds
(without alignments)
1403.047 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCSITKDKANPNNGQIQLE.....SSNENNADKIPGVRPCTFL 394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviris:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1200	59.9	271	2 Q49464	Q49464 mycoplasma
2	837	41.8	671	2 Q9L8D6	Q9L8D6 mycoplasma
3	831.5	41.5	702	2 Q49499	Q49499 mycoplasma
4	826	41.2	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	40.2	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	38.0	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	37.6	702	2 Q49498	Q49498 mycoplasma
8	746	37.2	645	2 Q9ZID1	Q9ZID1 mycoplasma
9	745	37.2	644	2 Q9ZHR9	Q9ZHR9 mycoplasma
10	744	37.1	650	2 Q49495	Q49495 mycoplasma
11	743	37.1	649	2 Q49497	Q49497 mycoplasma
12	713	35.6	647	2 Q49468	Q49468 mycoplasma
13	704	35.1	656	2 Q9KH15	Q9KH15 mycoplasma
14	690.5	34.5	703	2 O05122	O05122 mycoplasma
15	619.5	30.9	419	2 Q9L8D5	Q9L8D5 mycoplasma
16	575.5	28.7	386	2 Q49500	Q49500 mycoplasma

17	557	27.8	367	2 Q9XCG7	Q9XCG7 mycoplasma
18	189.5	9.5	320	2 Q49496	Q49496 mycoplasma
19	175	8.7	6713	16 Q99U54	Q99U54 staphylococ
20	175	8.7	6713	16 Q931R6	Q931R6 staphylococ
21	173	8.6	1302	2 Q49547	Q49547 mycoplasma
22	164.5	8.2	2481	16 Q99QR6	Q99QR6 staphylococ
23	163.5	8.2	3890	16 Q99U53	Q99U53 staphylococ
24	160.5	8.0	433	16 Q97TE6	Q97TE6 clostridium
25	160.5	8.0	682	2 Q48674	Q48674 lactococcus
26	159	7.9	933	2 Q53653	Q53653 staphylococ
27	158.5	7.9	1344	2 Q49545	Q49545 mycoplasma
28	156	7.8	807	5 Q96262	Q96262 plasmodium
29	155	7.7	1072	16 Q9CF64	Q9CF64 lactococcus
30	155	7.7	4888	16 Q9PQ08	Q9PQ08 ureaplasma
31	154.5	7.7	1365	2 Q49525	Q49525 mycoplasma
32	153.5	7.7	1795	2 Q9LCJ9	Q9LCJ9 staphylococ
33	153.5	7.7	2478	2 Q9RL69	Q9RL69 staphylococ
34	151.5	7.6	1237	16 Q9ZK57	Q9ZK57 helicobacte
35	151.5	7.6	2478	2 Q9LCH2	Q9LCH2 staphylococ
36	151	7.5	624	2 Q49548	Q49548 mycoplasma
37	150.5	7.5	604	5 Q26021	Q26021 plasmodium
38	150.5	7.5	1051	2 Q49524	Q49524 mycoplasma
39	149.5	7.5	962	2 Q49546	Q49546 mycoplasma
40	148	7.4	716	2 Q49526	Q49526 mycoplasma
41	147.5	7.4	655	5 Q26109	Q26109 plasmodium
42	146	7.3	854	2 Q9RLP9	Q9RLP9 mycoplasma
43	145	7.2	2399	16 Q9ZKS9	Q9ZKS9 helicobacte
44	144	7.2	1107	2 Q9F2D8	Q9F2D8 salmonella
45	143.5	7.2	751	3 Q94358	Q94358 schizosacch

ALIGNMENTS

RESULT 1

Q49464 ID Q49464 PRELIMINARY; PRT; 271 AA.
AC Q49464;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)
DE TM-1 (FRAGMENT).
GN TM-1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94025893; PubMed=8212828;
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,
RA Kamogawa K., Aoyama S., Iritani Y., Havashi Y.;
RT "Cloning and DNA sequence of a 29 kilodalton polypeptide gene of
RT Mycoplasma gallisepticum as a possible protective antigen.";
RL Vaccine 11:1061-1066(1993).
DR EMBL; S65869; AAB28343.2; -.
FT NON_TER 271
SQ SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;

Query Match 59.9%; Score 1200; DB 2; Length 271;
Best Local Similarity 97.1%; Pred. No. 2.2e-51;
Matches 238; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY	2	CMSITKDKANPNNGQIQLEAARMELTDLINAKAMTSLADYAKIEASLSAYSEATVN	61
DB	27	CMSITKDKANPNNGQIQLEAARMELTDLINAKAMTSLADYAKIEASLSAYSEATVN	86
QY	62	NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNVLVEAYKALKTTLLEQRATNLEGLS	121
DB	87	NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNVLVEAYKALKTTLLEQRATNLEGLA	146
QY	122	STAYNQIRNNLVLYNNKASSLITKTLTDLPLNGGTLLOSNEITTANKNINNTLSTINEQKTN	181

```
Db 147 STAYNQIRNNLVLDYNNASSLITKTLPLNGCMMLDSNEIITVARNINNTLSTINEQKTN 206
QY 182 ADALNSFIKVIQNNQSFYGTETNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
Db 207 ADALNSFIKVIQNNQSFYGTETNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266
QY 242 PSSRI 246
Db 267 PSSRI 271

RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
AC Q9L8D6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PMGA-LIKE PROTEIN 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.F., Branton S.L., Hanson L.A., Minlon F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel pmga-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAF29524.1; -.
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;

Query Match 41.8%; Score 837; DB 2; Length 671;
Best Local Similarity 43.4%; Pred. No. 2.5e-33;
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 5;

QY 9 DANI-----NNGQTOLEAARMELTDLINAKMTLASLDQYAKTEASLSAYSAETVNN 62
Db 55 DTNPGDGGMNSAQQLAAAKKELSDLLATQSSNLAKYADYTNLQNTLTAAYTTAKSTSD 114
QY 63 NLNATLEQLMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS 122
Db 115 NTSVLEQVKSATSTLQAAIDTAASSKTSFDEKNPELIKAYALKETLKNEETVLSGLTD 174
QY 123 TAYNOIRNNLVLDYNNASSLITKTLPLNGGTLTLDSEITANKNINNTLSTINEQKTN 182
Db 175 SNFATIKNTALYQSGKDFVKATLDVPVSGNA-POIADITKADKDIADAVSKLETWKINA 233
QY 183 DALNSFIKVIQNNQSFYGTETNANVQPSNYSFVAFSADVTP-----VNYKYARRTVNN 238
Db 234 NTLATSFVKELVKLNGKLTGIDT-TNNRQPGNYSFVGVSVNATNNNEIPNWFNAQRKVT 292
QY 239 GDEPSSRILA-----NTNITDVSIMYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAD 294
Db 293 SDNGRSLISTSDNSLTLEVSIMYSLSGAGTKYSLTFNYIGPSTGYLYFPYKLVKREGD 352
QY 295 ANNVGLOYLKNGNVQVEFA-----TSTSA-----NNTT-----ANPTPAVD 332
Db 353 ENNVALQYTLNLSGSAQEVNAPTVKTSVSDSSGDSNNQTESAAETWPTVSDLNPAPTVS 412
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 384
Db 413 DINIAKLTSLNLFKGSNTIEFSVPTPS--NKVAPMIGNIYLSNENAEKV 462

RESULT 3
Q49499 PRELIMINARY; PRT; 702 AA.
ID Q49499
AC Q49499;
```

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGAL 4 PROTEIN PRECURSOR.
GN PMGAL 4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=35010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62418.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT SIGNAL 1 25 POTENTIAL.
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;

Query Match 41.5%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 4.8e-33;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 9 DANPNN-----QTOLEAARMELTDLINAKMTLASLDQYAKTEASLSAYSAETVNN 62
Db 68 NTNPGNGGGTNAQAQQLAAAKKELSDLLATQSSNLSTYADYANTQNTLTAAYTTAKSTSD 127
QY 63 NLNATLEQLMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS 122
Db 128 NTSATLEQVKSATSTLQTAIDTAASSKTSFDEKNPELIKAYALKETLKWRNSLSGLTD 187
QY 123 TAYNOIRNNLVLDYNNASSLITKTLPLNGGTLTLDSEITANKNINNTLSTINEQKTN 182
Db 188 SNFATIKNTALYQSGKDIKTLTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKINA 246
QY 183 DALNSFIKVIQNNQSFYGTETNANVQPSNYSFVAFSADVTP-----VNYKYARTV 236
Db 247 TVLATSFVKELVKLNGKLTGIDT-TNNRQPGNYSFVGVSVDTTGGSDNARNPWSFAQRKV 305
QY 237 WNGD-----EPSSRILANTNITDVSIMYSLAGTNTKYQFSNYPSTGYLYFPYKLVK 291
Db 306 WTSNTDILSQPQPAEGENQQSQAPDVSIMYINLTGMAKYSLTFTNYGPGSTGYLYFPYKLVN 365
QY 292 AADANNVGLQYKLNNGNVQVEFATS-----TSANN-----TTANPTPAVD 332
Db 366 SSOSDKVALEYKLNESAVKTIIDFSQTSFVADATRENNRSTAAAPAGGSTEINPAPTLD 425
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNMKNVAPMIGNIYLSNENNADK 383
Db 426 DIKIAKLTSLNLFKGSNTIEFSVPTTAKGTSKVAPMIGNIYLSNENAEKV 477

RESULT 4
Q9XCG8 PRELIMINARY; PRT; 632 AA.
ID Q9XCG8
AC Q9XCG8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VLHA1 PRECURSOR (FRAGMENT).
GN VLHA1.
OS Mycoplasma imitans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=29560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4229;
```

RX MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Glew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans closely related to the pmga
family of Mycoplasma gallisepticum.";
RL Microbiology 145:2095-2103(1999).
RW EMBL; AF141940; AAD39483.1; -;
KW Signal.
FT SIGNAL. 1 1
FT NON_TER <1 17 POTENTIAL.
FT CHAIN 18 632 VLHAL.
SQ SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;

Query Match 41.2%; Score 826; DB 2; Length 632;
Best Local Similarity 45.3%; Pred. No. 9e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

QY 8 KDANPNNGOT-----QLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE 58
DB 34 KPNPDTGNTNPGGDTDAKLADAKASLNTLLGSOSTNVALYEDYAKIKDTLSSAYSAQ 93
QY 59 TVNNLNATLEOLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQATNLE 118
DB 94 TTANNANATLKDNDKAAQALQTAISDAANAaktefDKANGLVSAAYAKLTKLSETTND 153
QY 119 GLSSYAYNOIRNLDYLNKASSLTKLIDPLNGTGLDSNEITTANKNINNTLSTINEQ 178
DB 154 GLSDQNSYSAIKANLSLVNKAQDFTVTLDLP-TSGMIPKVDIEITSANTAITQAVSAIDSQ 212
QY 179 KTNADALSNSFKKVTIONEQSFVGTFTN---ANQPSNYSFVAFSADVTP----- 226
DB 213 KTNADATATTFKEKLDKSLKLI-PGTTAGQQAASQQPNYSFVGSNDVTTGRTGSQ 271
QY 227 --VNYKYARTVWNGDEPS---SRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTG 281
DB 272 DLPSWNNFAKKVWTEGLSAGOTQLVSSTPLTDVSWIYSLGAGSKYTLTFTYIGPSA 331
QY 282 YLYFFPKLVKAADANNVGLQYKLNNG-NVQOYEFATSSANTTANPTPAVDEIKVAKIV 340
DB 332 YLYFFPKLVQSSDSKNIGLQYKLNNSETLPTITFNGETN---DSGATPAIDDIKVKYT 387
QY 341 LSLGRLFGQNTIELSYPTGGNNKVPAMIGNIYLSNENNAADKI 384
DB 388 LSNLNFEGDNTIEFSVATEE---NKVAPMIGNIYLSNENNAADKI 428

RESULT 5
Q9KH13
ID Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ADHESIN PMGAL.4 (FRAGMENT).
GN PMGAL.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91415.1; -;
DR InterPro; IPR001986; EPSP_Synthese.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
FT NON_TER 584
SQ SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;

Query Match 40.2%; Score 805; DB 2; Length 584;
Best Local Similarity 42.4%; Pred. No. 7.6e-32;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

QY 11 NPNNNGOT-----QLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVNNLNA 66
DB 69 NPNSGNTTPEQOLAARKTLTDLTGENTNVALYADYAKIQSTLSTYATKATASENTSA 128
QY 67 TLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYN 126
DB 129 TLDNLRASASTTLQAAIDKAASNKRTFDSANQPLVTAYNQLKTTTQSKTTSLEGLSENKYS 188
QY 127 QTRNNLDYLNKASSLTKLIDPLNGTGLDSNEITTANKNINNTLSTNEOKTNADALS 186
DB 189 SIKNHLSKLFDAGSAIAAATLDP-TMGTVPEVMSVTKANEDIMTAVSKLTETWKTNAKRT- 246
QY 187 NSFIRKVIQNNQSFVGTFTNAN--VOPSNYSFVAFSADVTP-----VNYKYART 235
DB 247 NDFEKKPL--SKEKLVSTNDRAHQEPANWFSAGVSDLTGTGSGNSONLPNWFAQRK 304
QY 236 VVNGD---EPSSRIILANTNSITDVSWIYSLAGTNTKYQFSNYPSTGYLYFFPKLVKA 292
DB 305 VWTSEQQQTKTALVSSPVSATDVSWIYSLAGETKTYLTSFEYIGDPTAFLYFFPKLVKQ 364
QY 293 ADANNVGLQYKLNNGVQOYEF-----ATST-----S 319
DB 365 ADSSSVALQYSLNKTSSKLINFEPAKTMTNADQSENGVATTSTTEGRSSSEYLVADAVA 424
QY 320 ANNTTANPTPAVDEIKVAKIVSLGRLFGQNTIELSYPTGGNNKVPAMIGNIYLSNEN 379
DB 425 AVNNENMPTPTVSDINIAKVTLSGLTGFGEITFESVPT-----NKVAPMIGNIYLSNSG 479
QY 380 NADKI 384
DB 480 SOGKI 484

RESULT 6
Q9KH14
ID Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ADHESIN PMGAL.3.
GN PMGAL.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -;
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;

Query Match 38.0%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 1.2e-29;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

QY 9 DANPNNGOTQLEAARME-----LTDLINAKAMTSLASQDYAKIEASLSAYSEAEVYN 61
DB 63 DTNPGGGQGMNATNQELVNAKALSDLGESKTVELYADYAKIKADLTSAAYAKTTS 122
QY 62 NNLNATLEOLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
DB 123 DSTSTLDQVKATATLTQATINAAKDEKDFQDQNSQLLMAYKVLKDTLNKKEAIVNSLN 182

```
QY 122 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDNEITTANKNINNTL-----ST 174
Db 183 QEKYSAILSEINAASSTAEEIVKOTLPLVNG-----NLPVVAALNAENTKILEAIBEK 236
QY 175 INEOKTNADALSNSFIKKVIONNEQSFVGTFTTANVOPSNYSFVAFSADYTPVNYKYARR 234
Db 237 INSEKSNADLAFANYQLYKL---DRTKLMSGSSNTKPGNYSFVAYASDASPNWNAQR 293
QY 235 TVWNGD-----EPSSRILANTNSITDVSIIYSLAGTNKYQFSNYSNGPSTGYLYFPYKL 289
Db 294 TVWTFADSRWTSPLPNNLONSAPLTDVSWIYTLTGSGTGAKYTLTDPDYYPQOTGYLYFPYKL 353
QY 290 VKAADANNVGLQYKLNNGNVQVEF-----ATSTSANNTTANPTPAV 331
Db 354 VKTSD--KVLQYKLNQADPAVQFSAAATASAPATDGRQESAETATANEKVNPMPSV 411
QY 332 DEIKVAKIVLSGLRFGQNTIELSYPTGEGNMKNKVPAMIGNIYLSNENNADK 383
Db 412 NTINAVKTVLSNLKFGSNTIEFSVPMDDNNKVPAMIGNIYITSSNDANK 463
RESULT 7
Q49498 PRELIMINARY; PRT; 702 AA.
AC Q49498;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE PNGA1.3 PROTEIN PRECURSOR.
GN PNGA1.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95010739; PubMed=7925999;
RX Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multi-gene family which encodes the major cell
RL surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62417.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 702 AA; 75537 MW; 273E8915FEE57B9F CRC64;

Query Match 37.6%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 2.8e-29;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 11 NPNGOT----QLEAARMELTDLINAKMTLASIQYAKIEASLSAYSEATVNNLNA 66
Db 69 NPNGNTPPQQLAARKTLTDLGENTVVALYADYAKIQSTLSTAYTAKTASENTSA 128
QY 67 TLBOLKMAKTNLBSAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLSPAYN 126
Db 129 TLENLSASTTLQAAIDKAANDKRVDSVNOPLVAAYNLTKLTKSTTSLEGLSENKY 188
QY 127 QIRNNVLDLYNKASSLITKTLPLNG--GTLTLDNEITTANKNINNTLS--TNEOKTNA 182
Db 189 GIKNHLKSLFDTSATKATKTLDTSGERPTLEKVN---ANNGIKMAISPESLKKWKGNA 245
QY 183 DALSNSFIKKVIONNEQSFVGTFTTANVOPSNYSFVAFSADYTP-----VNYKYARTVW 237
Db 246 DKF-NEFEKNPLSEKELKSTSDTAHQEQAPANSPAYSDLTLSNQNLPNWFNAQKVV 304
QY 238 NGD--BPSSRILANTN-SITDVSIIYSLAGTNKYQFSNYSNGPSTGYLYFPYKLKAAAD 294
Db 305 TSENQOPGKTALVSSPVASDEIVSWIYSLAGEGTYTLTTFEYGPDAFLPYKLVKAAAD 364
```

```
QY 295 ANNYGLQYKLNNGNVQVEF-----ATST-----SAN 321
Db 365 SSSVALQYSLNKTSSKLNFPKPAETVSTNTDQSENEVATTSTTEARSYKVLVADAATS 424
QY 322 NTTANPTPAVDIEIKVAKIVLSGLRFGQNTIELSYPTGEGNMKNKVPAMIGNIYLSNENNA 381
Db 425 NNEMHPTPTSDINIAKIVTSLGTLFGGNTIEFSVPEG-----KVAPMIGNIYLSNSESQ 479
QY 382 DK1 384
Db 480 VK1 482
RESULT 8
Q9ZIDI PRELIMINARY; PRT; 645 AA.
AC Q9ZIDI;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE M9 PROTEIN.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31; ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family.";
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;

Query Match 37.2%; Score 746; DB 2; Length 645;
Best Local Similarity 42.3%; Pred. No. 6.2e-29;
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;

QY 9 DAPNNGQ-----TOLEAARMELTDLINAKMTLASIQYAKIEASLSAYSEATVN 61
Db 44 DTNPGDQGGMMNAASQELAAARMGLTTFVDSKAKNLGLYDYKKTQDTLTAKYDAKTVL 103
QY 62 NNLNATLEQLKMAKTNLBSAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 104 DNSSTTQNLNEAKRLETAIRTAATSKQTFDEQRAELVKYEEKTTLSTNETATLAPYA 163
QY 122 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDNEITTANKNINNTL--STINEOK 179
Db 164 AAQYAGIKHLSGLYDAGKAITTKTLEPVEGDP-LTADVVMANTKIVEAIKDEVLPQK 222
QY 180 TNADALSNSFIKKVIONNEQSFVGTFTTANVOPSNYSFVAFSADYTPV-----NKKY 231
Db 223 ENATKLADSEFVKVLVKEKITGVEEAHN-KAQPANYSFVGVSDITCTVTGTQTSIPNDY 281
QY 232 ARTVW-NGDEPSSRILANT-----NSITDVSIIYSLAGTNKYQFSNYSNGPSTGY 282
Db 282 AQRTIFTNGDEP--RSISNTPADGQTMVQPLSNYSWISYLAGTKYKTLFTTYGPGSTGY 339
QY 283 LYFPYKLVKAAADANNVGLQYKLNNGNVQVEFATSTSA-----NNTTANPTPAVDELKVA 337
Db 340 LYFPYKLVNTSDQMKLGLEYKLN-----ATEPSAIFGSEQTNGTKTPTVNDINVA 391
QY 338 KIVLSGLRFGQNTIELSYPTGEGNMKNKVPAMIGNIYLSNENNADK1 384
Db 392 KVTLANLKGSKNIEFSVPA-----EKVSPMIGNIYLSNENNNKI 433
RESULT 9
```



```
Q92HR9          PRELIMINARY;          PRT;      644 AA.
ID Q92HR9
AC Q92HR9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (MG) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RL Infect. Immun. 66:5570-5575(1998).
RW EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

Query Match      37.2%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 7e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

QY 9 DAMPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 61
Db 44 DTNPGDGGGMNNAASQELAAARMGLTTFVDSKAKNLGLVYDYKKTQNTLTTRAYDAKTVL 103
QY 62 NNLNATLEQIKWAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQKATNLEGLS 121
Db 104 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDQHAELVKVYEELKTTLSNETATLAPYA 163
QY 122 STAYNOIRNLVDLYNKASLTKTLDPLNGTLLDSNEITTANKNNLTL--STNEOK 179
Db 164 AQAQAGIKMHLGSLYDAGAKITTKTLEPVEGDP-LTASAVMANTRKIVEAIKDEVLPQK 222
QY 180 TNADALSNSFIKKVQNNQSFVGTFTNANVQPSNYSFVAFSADTPV-----NRYK 231
Db 223 ENATKLADFVKQVLYKEKITGVEEAHNS-QPANTSFVGSVDITGTANGQTSIPNNWY 281
QY 232 ARTVW-NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 282
Db 282 AORTIFTNGDEP--RSVSNTPVDGQMAQPLSNVSWIYSLAGTGAKYTLFTFYGPSTGY 339
QY 283 LYFPYKLVKAADANNVGLQYKLNNGNVQVEFATSTSA-----NNTTANPTPAVDIKVA 337
Db 340 LYFPYKLVNTSDQVKLGLEYKLNDA-----ATKPSAITFGSEQTPMNGKTPVTNDINVA 391
QY 338 KIVLSGLRFQNTIELSVPTGEGNMKNKVPMTGNITVLSNENNAADKI 384
Db 392 KYTLANLNFSGNKIEFSVP-----VEKVSFPMIGNMYLSSSPNNWNKI 433

RESULT 10
Q49495          PRELIMINARY;          PRT;      650 AA.
ID Q49495
AC Q49495; O08060;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HAEMAGGLUTININ PRECURSOR.
DE PMGAL.1
GN PMGAL.1
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50152.1; -.
KW Signal.
FT SIGNAL.
FT CHAIN 26 650 HAEMAGGLUTININ.
FT CHAIN 26 650 HAEMAGGLUTININ.
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;

Query Match      37.1%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 7.9e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

QY 11 NPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVNNN 63
Db 51 NPGDGGGMNNAASQELAAARMGLTTFVDSKAKNLGLVYDYKKTQNTLTTRAYDAKTVL 110
QY 64 LNAATLQKMAKNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQKATNLEGLSST 123
Db 111 SSSSTQNLNEAKTRLETAIRTAATSKQTFDQHAELVKVYEELKTTLSNETATLAPYA 170
QY 124 AYNQIRNLVDLYNKASLTKTLDPLNGTLLDSNEITTANKNNLTL--STNEOKTN 181
Db 171 QYAGIKMHLGSLYDAGAKITTKTLEPVEGDP-LTACAVTMANTKIVEAIKDEVLPNKEN 229
QY 182 ADALSNSFIKKVQNNQSFVGTFTNANVQPSNYSFVAFSADTPV-----NRYKAR 233
Db 230 ATKLADSFVKQVLYKEKITGVEEAHN-KAQPANYSFVGSVDITGTANGQTSIPNNWY 288
QY 234 RTVW-NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 284
Db 289 RIIFTNGDEP--RSVSNTPVDGQMAQPLSNVSWIYSLAGTGAKYTLFTFYGPSTGY 346
QY 285 FPKYKLVKAADANNVGLQYKLNNGNVQVEFATSTANNTAN-PTPAVDIKVAKIVLSG 343
Db 347 FPKYKLVNTSDQVKLGLEYKLNDA-----TEPSAITFGNEQTMNGKTPVTNDINVAKTVL 402
QY 344 LRFQNTIELSVPTGEGNMKNKVPMTGNITVLSNENNAADKI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 438

RESULT 11
Q49497          PRELIMINARY;          PRT;      649 AA.
ID Q49497
AC Q49497;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
DE PMGAL.2
GN PMGAL.2
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN [1]
RP SEQUENCE FROM N.A.
```

RL EMBL: 352:347-352(1994).

DR EMBL: L28424; AAA62416.1; -.

KW Signal.

FT SIGNAL 1 25 POTENTIAL.

SQ SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 37.18; Score 743; DB 2; Length 649;

Best Local Similarity 42.3%; Pred. No. 8.9e-29;

Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

QY 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEFTN 61

DB 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVL 107

QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121

DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167

QY 122 STAYNOIRNNLVLYNKASSLIITKTLPLNGGTLSDNEITTANKNINNTL--STINEQK 179

DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLPQK 226

QY 180 TNADALSNSFKKVIQNNQSFVGTETNANVPQSNYSFVAFSADVTPV-----NKY 231

DB 227 ENATKADSFVKVLVKEKITGVVEAHN-KAQPANYSFVGSVDITGTTGQTSIPNDY 285

QY 232 ARRTVW-NGDEPSSRIILANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGY 282

DB 286 AQRITFTNSDEP--RSISNTPADQQTMAQPLSNVSWIYSLAGTCAKYTLEFTYYPSTGY 343

QY 283 LYFPYKLVKAADANNVGLQKLNNGVQVEFATSTSA-----NNTTANTPVADEIKVA 337

DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395

QY 338 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNNENNAKDI 384

DB 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 12

Q49468

ID Q49468 PRELIMINARY; PRT; 647 AA.

AC Q49468; Q53303;

DT 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)

DE HEMAGGLUTININ HOMOLOG PRECURSOR.

GN PMGAL.2.

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

OC Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-93162830; PubMed-8432610;

RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;

RT "Molecular cloning of a member of the gene family that encodes pmGA, a

RT hemagglutinin of Mycoplasma gallisepticum.";

RL Infect. Immun. 61:903-909(1993).

DR EMBL; M83178; AAA02996.1; -.

DR EMBL; S55216; AAB25397.2; -.

KW Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.

SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match

Best Local Similarity 35.6%; Score 713; DB 2; Length 647;

Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

QY 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEFTN 61

~~*

DB 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVL 107

QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121

DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167

QY 122 STAYNOIRNNLVLYNKASSLIITKTLPLNGGTLSDNEITTANKNINNTL--STINEQK 179

DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLPQK 226

QY 180 TN-----ADALSNSFKKVIQNNQSFVGTETNANVPQSNYSFVAFSADVTPV----- 227

DB 227 ENATKADLSLSSIVKKITGVVE-----AHNKAQPANYSFVGYKRWYTELLDKQVFP 279

QY 228 NYKARTVW-NGDEPSSRIILANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGY 278

DB 280 NWDYAQRTFTNSDEP--RSISNTPADQQTMAQPLSNVSWIYSLAGTCAKYTLEFTYYPG 337

QY 279 STGYLYFPYKLVKAADANNVGLQKLNNGVQVEFATSTSA-----NNTTANTPVADE 333

DB 338 STGYLYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVND 389

QY 334 IKYAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNNENNAKDI 384

DB 390 INVAKVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 435

RESULT 13

Q9KH15

ID Q9KH15 PRELIMINARY; PRT; 656 AA.

AC Q9KH15;

DT 01-OCT-2000 (TEMBLrel. 15, Created)

DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)

DE ADHESIN PMGAL.2.

GN PMGAL.2.

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

OC Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-HS;

RA Shen Q.C., Bi D.R., Weng C.J.;

RT "Sequence analysis of the pmGA multigene family of Mycoplasma

RT gallisepticum strain HS.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF275312; AAF91413.1; -.

SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;

Query Match 35.1%; Score 704; DB 2; Length 656;

Best Local Similarity 41.0%; Pred. No. 6.9e-27;

Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

QY 9 DANPNNGQ-----QLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEFTN 61

DB 63 DINGPGGQNMDSAAQELTAARTALTSLLASKNANVEMYSYAKIQNTLAAVTTAQTQS 122

QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121

DB 123 QNSSATLEQVKNATSLQTAINTANSNKQKQDQDHSNLLMSYKKNMATLAKKETAVMTLK 182

QY 122 STAYNOIRNNLVLYNKASSLIITKTLPLNGGTLSDNEITTANKNINNTL--TINEQK 179

DB 183 DPYSAILDQINGVSSKGELVQHTLDPVS-GIVPRANTITEETKEEIVISEKTLODQK 241

QY 180 TNADALSNS--SFI--KKVIQNNQSFVGTETNANVPQSNYSFVAFSADVTPV----- 227

DB 242 NNADQFANYQSFTLDKTKLENVEDA-----KKMGQPANYSFVGYSDVTGTSGQETIP 295

QY 228 NYKARTVWNGDEPSSRIILANTNS-----ITDVSWIYSLAGTNTKYQFSNYPSTGY 279

```
Db 296 NWAFQRAITSGNQTKYTATTGGEDQSTAKPLSDVSWIYSLAGTGAKYLFTEYIGPS 355
QY 280 TGYLYPPYKLVKAADANNVGLQYKLN-NGVQVFEFATSTSANNTANPTTANPPAYDEIKVAK 338
Db 356 TGYLYPPYKLVKAND--DVGLOQKLSNETLTPFIIFEGT-----TTNGPAATVENINVAK 409
QY 339 IVLSGLRFGQNTIELSPTEGEGNNKVPAMIGNIYSSNENNAK 383
Db 410 VRLTGLAFGKNTIEFSVP-----MSKVAPMIGNIYSSDTETNK 449

RESULT 14
O05122 PRELIMINARY; PRT; 703 AA.
AC O05122;
DT 01-JUL-1997 (TREMREL. 04, Created)
DT 01-JUL-1997 (TREMREL. 04, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE HAEMAGGLUTININ.
GN PMGAL.9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50154.1; -.
DR InterPro: IPR002819; HD.
SQ SEQUENCE 703 AA; 75742 MW; 310869BE9F73CBC5 CRC64;

Query Match 34.5%; Score 690.5; DB 2; Length 703;
Best Local Similarity 35.0%; Pred. No. 3.4e-26;
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;

QY 2 CMSTTKKDANP-----NNGQT-----OLEARMELTDLINAK 33
Db 26 CTSATIPLTNPPEKPPDPMPNPPSGMGNGTNGMDTAQELASAKAALTITNRE 85

QY 34 AMTLASLDYAKIEASLSAYSEATVNNLNATLEQKMAKTNLESAINQANTDKTTFD 93
Db 86 SEKVGLVDYAKIKADLTSAITVAKTSDSTSTLVQVKTATSKLTQTAIDKAASDKQFE 145

QY 94 NEHPNLVEAYKALKTLEQR-ANLEGLSTAYNQIRNNLVLYNKASSLITKTLDPLNG 152
Db 146 QDRHKDLMPYSELKTLTSQKNATVL--LNQPKYSAILNKINSIYAQGEVVIITLDVPS- 202

QY 153 GTLLDSNEITANKNNITLS--TINEOKTNADALSNSFKKVIQNNQSFVGTFTNANV 210
Db 203 GAIPTRASITKVNDEINKAISLENOLPKPKDNADAFANYQFFKL---DKTIMGMSTNMK 259

QY 211 QPNYSFVAFSADVTPV-----NYKARTVWNGDEPSSRILANTNS----- 252
Db 260 QPNYSFVGYSVGTGQSQGTIPNNFAQRIWSSGAPRAPLASQETFPQAEPTPMSA 319

QY 253 -----TTDVSWIYSLA 263
Db 320 POGVEPAQQQDSSPKQASSETQEVSPTPAAEQAAQADTEQPATSOQTPLTLDVSWIYSL 379

QY 264 GTNTKYQFSNYPGSPSTGYLYFFPKLVKAADANNVGLQYKLNNGVQVFEFATSTSANNT 323
```

```
Db 380 GTDVKTTFTFNTFPGPMAYLYFFPKLVKSDD--SVGLQYKLNNNNPNVALNFGSETNAN-- 435
QY 324 TANPTDEIKVAKIVLSGLRFGQNTIELSPTEGEGNNKVPAMIGNIYSSNENNAK 380
Db 436 --GPAASVDNINVAKNVLANLNFEGNTIEFSVP-----MKNKVPAMIGNIYSSDVAN 485

RESULT 15
Q9L8D5 PRELIMINARY; PRT; 419 AA.
AC Q9L8D5;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-OCT-2000 (TREMREL. 15, Last annotation update)
DE PMGA-LIKE PROTEIN 9.3 (FRAGMENT).
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAF29525.1; -.
FT NON_TER. 419
SQ SEQUENCE 419 AA; 45072 MW; FE5EBE37F2DB3B0C CRC64;

Query Match 30.9%; Score 619.5; DB 2; Length 419;
Best Local Similarity 40.1%; Pred. No. 5.4e-23;
Matches 149; Conservative 58; Mismatches 122; Indels 43; Gaps 11;

QY 9 DANPNNGQT-----OLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 61
Db 61 DINPGGQNMDSAAQELTAARTALTSLLASKNANIEMYSYAKIQTIONTLIAAYTTAEQTS 120

QY 62 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTLEQRTNLEGLS 121
Db 121 QNSATLEQVKNATSALTQTAINTANSNKQKFDQDHSNLLMSYKNNMLATLAKETTVMTLK 180

QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGGLTLLDSNEITANKNNITLS--TINEQK 179
Db 181 DPYSAILDQINGSCKEELVQHTLDVPS-GIVPAANTITEITKIEEVISEKTLQDQK 239

QY 180 TNADALS- -SFT--KKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 227
Db 240 NNADQFDNYQSFTLDKTKLENVEDA-----KKMGQPANYSFVGYSVDVTGTSQGETTP 293

QY 228 NYKARTVWNGDEPSSRILANTNS-----ITDVSWIYSLAGTNKYQFSNYPGSP 279
Db 294 NWAFQRAITSGNQTKYTATTGGEDQSTAKPLSDVSWIYSLAGTGAKYLFTEYIGPS 353

QY 280 TGYLYPPYKLVKAADANNVGLQYKLN-NGVQVFEFATSTSANNTANPTTANPPAYDEIKVAK 338
Db 354 TGYLYPPYKLVKAND--DVGLOQKLSNETLTPFIIFEGT-----TTNGPAATVENINVAK 407

QY 339 IVLSGLRFGQNT 350
Db 408 VRLTGLAFGKNT 419

Search completed: June 12, 2002, 10:50:42
Job time: 205 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:51:05 ; Search time 17.44 Seconds
(without alignments)
874.742 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSITKDKANPNNGQIQLE.....SSNNENADKIPGRRPGTFL 394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	7.7	682	1 NISP_LACLA	Q07596 lactococcus
2	143	7.1	857	1 AR56_CANAL	P78586 c arg5.6 pr
3	138.5	6.9	1314	1 SWIL_YEAST	P09547 saccharomyc
4	137	6.8	719	1 YM41_YEAST	Q03213 saccharomyc
5	135.5	6.8	1085	1 CUT7_SCHPO	P24339 schizosacch
6	132	6.6	1271	1 Y338_MYCGE	P47580 mycoplasma
7	131.5	6.6	1164	1 BAG_STRAG	P27951 streptococ
8	129	6.4	1251	1 RBP2_PLAVB	Q00799 plasmodium
9	128.5	6.4	2334	1 WAPA_BACSU	Q07833 bacillus su
10	128.5	6.4	3712	1 LMA_DROME	Q00174 drosophila
11	128	6.4	995	1 YI09_YEAST	P40442 saccharomyc
12	127	6.3	956	1 YEF3_YEAST	P32618 saccharomyc
13	126.5	6.3	444	1 PST1_YEAST	Q12355 saccharomyc
14	126.5	6.3	1630	1 MSP1_PLAFW	P04932 plasmodium
15	126.5	6.3	1839	1 MSP1_PLAFW	P04933 plasmodium
16	125.5	6.3	1858	1 P3K2_DICDI	P54674 dictyosteli
17	124.5	6.2	821	1 LIK5_CAEEL	P45970 caenorhabdi
18	124.5	6.2	1010	1 YK11_CAEEL	P34278 caenorhabdi
19	124.5	6.2	1018	1 FNBA_STAUA	P14738 staphylococ
20	124	6.2	749	1 MAD1_YEAST	P40957 saccharomyc
21	124	6.2	750	1 YK57_YEAST	P34331 saccharomyc
22	124	6.2	823	1 NSPL_YEAST	P14907 saccharomyc
23	124	6.2	866	1 MYSP_SCHJA	Q05870 schistosoma
24	124	6.2	1117	1 YN96_YEAST	P53753 saccharomyc
25	124	6.2	1159	1 N124_SCHPO	Q09904 schizosacch
26	124	6.2	1433	1 CAP8_YEAST	P39113 saccharomyc
27	123.5	6.2	537	1 ARP_PLAFA	P04931 plasmodium
28	123.5	6.2	1902	1 P2P_LACIC	P15293 lactococcus
29	123	6.1	1260	1 ALS1_CANAL	P46590 candida alb
30	123	6.1	1637	1 MRSP_STAUA	P80544 staphylococ
31	123	6.1	1790	1 US01_YEAST	P25386 saccharomyc
32	122.5	6.1	507	1 FLIC_SALON	Q06974 salmonella
33	122	6.1	650	1 PDR4_YEAST	P19880 saccharomyc

ALIGNMENTS

RESULT 1

ID	NISP_LACLA	STANDARD;	PRT;	682 AA.
AC	Q07596;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Nisin leader peptide processing serine protease NISP precursor			
DE	(EC 3.4.21.-)			
GN	NISP.			
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Lactococcus.			
OX	NCBI_TaxID=1360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NIZO R5;			
RX	MEDLINE=93239683; PubMed=8478324;			
RA	van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,			
RA	Kuipers O.P., de Vos W.M.;			
RT	"Characterization of the Lactococcus lactis nisin A operon genes			
RT	nisp, encoding a subtilisin-like serine protease involved in			
RT	precursor processing, and nispR, encoding a regulatory protein			
RT	involved in nisin biosynthesis."			
RL	J. Bacteriol. 175:2578-2588(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=6F3;			
RX	MEDLINE=94213458; PubMed=8161176;			
RA	Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,			
RA	Hammelmann M., Enian K.-D.;			
RT	"Regulation of nisin biosynthesis and immunity in Lactococcus lactis			
RT	6F3."			
RL	Appl. Environ. Microbiol. 60:814-825(1994).			
RN	[3]			
RP	3D-STRUCTURE MODELING.			
RX	MEDLINE=95357326; PubMed=7630881;			
RA	Siezen R.J., Rollena H.S., Kuipers O.P., de Vos W.M.;			
RT	"Homology modelling of the Lactococcus lactis leader peptidase Nisp			
RT	and its interaction with the precursor of the lantibiotic nisin."			
RL	Protein Eng. 8:117-125(1995).			
CC	-1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.			
CC	-1- PATHWAY: LAST STEP OF NISIN BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL			
CC	(POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE			
CC	SUBTILASE FAMILY.			
CC	-1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS			
CC	IN THE REGION OF THE MEMBRANE ANCHOR.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>)

Q9PIY2 chlamydia m
P38272 saccharomyc
Q04893 saccharomyc
Q54001 salmonella
Q09625 caenorhabdi
P49022 lactococcus
P41508 mycoplasma
P28742 saccharomyc
Q02470 lactobacilli
Q02945 equine rota
P25146 listeria mo
P11657 streptococ

34 122 6.1 1672 1 PMPB_CHLMU
35 121.5 6.1 425 1 YBY0_YEAST
36 121.5 6.1 1140 1 YBY0_YEAST
37 121 6.0 491 1 TOLC_SALEN
38 121 6.0 796 1 YS8A_CAEEL
39 121 6.0 901 1 PIP_LACLA
40 121 6.0 979 1 P115_WYCHR
41 121 6.0 1111 1 KIP1_YEAST
42 121 6.0 1902 1 P2P_LACPA
43 120.5 6.0 776 1 VP4_ROTPEH
44 120.5 6.0 800 1 INLA_LISMO
45 120.5 6.0 1565 1 PAC_STRMU

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; L11061; AAA25200.1; -.
DR EMBL; X76884; CAA54210.1; -.
DR HSP; P29600; IGC1.
DR MEROPS; S08.059; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
DR PROSITE; PS00344; GRAM_POS_ANCHORING; 1.
KW Hydrolase; Serine protease; Cell wall; Signal; Transmembrane; Zymogen.
FT SIGNAL 1 22
FT PROPEP 23 195
FT CHAIN 196 682
FT FT
FT DOMAIN 196 651
FT TRANSMEM 652 679
FT DOMAIN 680 682
FT ACT_SITE 259 259
FT ACT_SITE 306 306
FT ACT_SITE 512 512
FT DOMAIN 652 657
FT FT
FT CONFLICT 500 500 A -> T (IN REF. 2).
FT SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Query Match
Best Local Similarity 7.7%; Score 154.5; DB 1; Length 682;
Matches 75; Conservative 21.0%; Pred. No. 0.11;
Matches 75; Indels 91; Gaps 13;

QY 58 ETVNNLNATLQOLMAKNTLESAINQANTKTTFDNEHPNLVEAYKALKTKTLEORATNL 117
Db 38 ELIHNSNALSSTEGSTDSINLGAQSPAVKSTTRTE-----LDVTGAATLLQTSVQK 93

QY 118 EGLSSPAYQIRNVLNLDKASSLITKTLPLNGTGLDLSNEITANKNINNTLSTINE 177
Db 94 EMKVSLOETQVASE-----FSKRDVYNKEAVPSKDELLEQSEVVSTSSIQKN-KILDN 148

QY 178 QKTNADALNSFTKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVPVNYKYARRV- 236
Db 149 KKRANFVTSPLIKEKPSKSDASGVIDNS-----ASPLSYRKAKEVVS 193

QY 237 -----WNGDEPSSRLANTNSITDVSWIYSLAGTNTKYQFSFSGNYPGS 279
Db 194 LRQLNKNQKVEAQLLISNSSEKASVYTNSHDFWDYQW-----DMKY---VTNNGES 243

QY 280 TGYLYFPYKLVKAADANNVGLQYKLNNGVQVFEATSTAN-----NT 323
Db 244 YA-LYQPSKKI-----SVGI-----IDSGIMEHPDLSNLSGNYFKNLVPKGFGFDNEPDE 293

QY 324 TAPPTPAVDIEKIVLGLRQNTIELSVPTGSENNKVPALMG-NIYLSSEN 379
Db 294 TGNPSPDIVD-----KMGHGTVEAGQITANGNIIIGVAPGITVNIYRVFGEN 338

RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE AR56.6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN AR56.6.

```

```

OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OC NCBI_taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans";
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -1- PATHWAY: SECOND AND THIRD STEPS IN ARGinine BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; X98880; CAA67383.1; -.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR001048; Aakkinase.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakkinase; 1.
DR Pfam; PF01118; Semialdehyde_dhc; 1.
DR Pfam; PF02774; Semialdehyde_dhc; 1.
DR PRODOM; PD003765; AGPR_act_site; 1.
DR PROSITE; PS01224; ARGC; 1.
DR Oxidoreductase; Transferrase; Kinase; Arginine biosynthesis; NADP;
DR Mitochondrion; Multifunctional enzyme; Trans peptidase.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT ACT_SITE 669 669 BY SIMILARITY.
FT SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C4C71 CRC64;

Query Match
Best Local Similarity 7.1%; Score 143; DB 1; Length 857;
Matches 99; Conservative 23.3%; Pred. No. 0.64;
Matches 99; Indels 134; Gaps 25;

QY 46 IEASLSSAY-----SEATVNN---NLNATL-----EQLKMAKNTLESAINQANTD 88
Db 208 VEAAINSGLYPLTSLAETSSQLLNVDVAAAGELAREFEPLKIYLVNKGRIINGTG 267

QY 89 K-----TTDNEHPNLV-EAYKALKTKTLEOR-----ATNLEGLSSPAYQIRNVLNLDYLNK 138
Db 268 EKVSAINDEEYEDLLKESWYKYGKLIKKEIHDLLQHLPRSSVA-----IIDVNDL 320

QY 139 ASSLITKTLDPNGGTLSDSNEITTANKNIN-NTLSTINEOKTNADALNSFIKKVIONN 197
Db 321 QKELFTDS-----GAGTL-----IRRGYRLINRNSLRDFG-----NPDLLRNALLR 361

QY 198 EQSFVGTFTNANVQPSNYSFVAFSADVPVNYKYARRVWNGDEPSSRLANT----- 250
Db 362 -----DPEIKTKVSVASYLKFDSVQFAS-----YGDPE-LEVLAIVQNDKI 405

QY 251 -----NSITD-----VSWIYSLAGTNTKYQFSFSGNYPG-PSTGYL 283
Db 406 PKLDFLSSKGTGLNNTDNIENAIKKDYSQLCWVYVNDANLWPYFSKSDGSFARNGQI 465

```



```

QY 252 SIT--DYSWIYSLAGTNTKYQFSFNSYCPSTGYLYFFPKLVKADANNVGLOYKLNNGNV 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 533 NFSKDNVSFSDVDFGNSSEFR-SWANRN--ITLKLYTALTITMLENGTSN-----NNGQK 583
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 310 QQVEFATSTSANNTNAPT 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 DVCDLAKKLLKNNTNLSET 602
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
BAG_STRAG
ID -BAG_STRAG STANDARD; PRT; 1164 AA.
AC P27951;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IGA FC receptor precursor (Beta antigen) (B antigen).
DE GN BAG.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-48.
RC STRAIN=LA239;
RC MEDLINE=91312121; PubMed=1857207;
RA Jarlstrom P.G., Chhatwal G.S., Timmis K.N.;
RT "The IGA-binding beta antigen of the c protein complex of Group B
RT streptococci: sequence determination of its gene and detection of two
RT binding regions.";
RL Mol. Microbiol. 5:843-849(1991).
RN [2]
RP IDENTIFICATION OF IG-LIKE DOMAIN.
RX MEDLINE=97035265; PubMed=8880921;
RA Bateman A., Eddy S.R., Clothia C.;
RT "Members of the immunoglobulin superfamily in bacteria.*";
RL Protein Sci. 5:1939-1942(1996).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59771; CAA42442.1; -.
CC PIR: S15330; FCSOAG.
CC InterPro: IPR001899; Gram_pos_anchor.
CC InterPro: IPR003599; Ig.
CC Pfam: PF00746; Gram_pos_anchor; 1.
CC SMART: SM00409; IG; 1.
CC PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
CC Cell wall; Transmembrane; Receptor; Repeat; Signal;
CC Immunoglobulin domain.
CC SIGNAL 1 37
CC FT CHAIN 38 1164 IGA FC RECEPTOR.
CC FT DOMAIN 38 1131 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 1141 1159 MEMBRANE ANCHOR (POTENTIAL).
CC FT DOMAIN 1160 1164 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 434 534 IG-LIKE DOMAIN.
CC FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).
CC FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
CC FT DOMAIN 827 945 PRO-RICH REPEATS.
CC FT DOMAIN 1131 1137 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
CC PROTEINS.
CC SEQUENCE 1164 AA; 131051 MW; 65DE94AF720A5474 CRC64;
SQ

```

		Matches	89;	Conservative	70;	Mismatches	140;	Indels	122;	Gaps	20;
QY	16	OTOLEAARMELDLNNAKAMTLASQDYAKYKAEASLSAYSAE-----TVNNLNAT	67								
Db	424	ETNLSVRNLSITNIK-----QGGEKI-----YSRAKDIMQIKATSENTAETK	469								
QY	68	LEOLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEORATNLEGLSTAYNQ	127								
Db	470	LEKVKDDQSNVYNYLNQITTR-----NLIVTEKNRLNGIDSTITNIEG-----ALKE	517								
QY	128	IRNNL-----VDLYNKASSLITKTLPLDPLNGGTLTLLDSNEIT--TANKN	167								
Db	518	SKNGYIGLEKLEEGKKNRKLKVDITKKS---INSTVG--NFSSLFNNFDLNOYDFNKN	572								
QY	168	INTTLSTINEQKTADALNSNF---IKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADV	224								
Db	573	INDYENKMG-----IYNEPEGLNKISEN-----LRNASENTSDYN---SAKT	613								
QY	225	TPVNYKYARTVWNGDEPSSRLANTNSITDYSWIY---SLAGTN---TKYQFSFNYG	277								
Db	614	LRLEAQEKVYLNLLKKEEANKYLRDVKVFIFNMKESLDKINEMIKKEQLFPVNEGH	673								
QY	278	PSGILYFPYKLVKAADANNVGLQYKLNNGVQOQVEFATSTANN-----TTA--	325								
Db	674	GNVKQLVENIK--ELYDENNLSDILQAOTGKNEEIQKITHSTLKNKAKTILGHVDTSAKY	731								
QY	326	-----NPTPAVDE-----IKVAK-----IVLSGLRFGONTIELSVPTGEGNMKV	365								
Db	732	VGIKITPELALTLLGDAKLKTAQELKFESKNVNVLETENMKNSTNELDVHKNQIDAYKY	791								
QY	366	A 366									
Db	792	A 792									

RESULT 9	
WAPA_BACSU	STANDARD; PRT; 2334 AA.
ID	WAPA_BACSU
AC	Q07833; 1994 (Rel. 30, Created)
DT	01-OCT-1994 (Rel. 30, Last sequence update)
DT	01-OCT-1994 (Rel. 30, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Wall-associated protein precursor.
GN	WAPA OR N17G.
OS	Bacillus subtilis.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;
OC	Bacillus/Staphylococcus group; Bacillus.
OX	NCBI_TaxID=1423;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=168;
RX	MEDLINE=93302506; PubMed=8316082;
RA	Foster S.J.;
RT	"Molecular analysis of three major wall-associated proteins of
RT	Bacillus subtilis 168: evidence for processing of the product of a
RT	gene encoding a 258 kDa precursor two-domain ligand-binding
RT	protein.";
RL	Mol. Microbiol. 8:299-310(1993).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=168 / BGSC1AL;
RX	MEDLINE=95219088; PubMed=7704263;
RA	Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT	"Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT	genome containing the hut and wapa loci.";
RL	Microbiology 141:337-343(1995).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=168 / BGSC1AL;
RX	MEDLINE=97124196; PubMed=8969509;
RA	Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA	Miwa Y., Fujita Y.;
RT	"Sequencing of a 65 kb region of the Bacillus subtilis genome

Query Match		6.6%;	Score 131.5;	DB 1;	Length 1164;
Best Local Similarity		22.8%;	Pred. No. 4;		
Matches 84;		Conservative 51;	Mismatches 157;	Indels 77;	Gaps 15;
QY	7	KKDANPNNGTQLEA-ARMELTDLINA-KAMTLASIQDYAKYKAEASLSAYSE-----AET	59		
Db	230	RKQQAQADKKEDAEVKVREELKGLFSKAGLDQEOIEQHVKKETSSSENTQKVDEHYANS	289		
QY	60	VNNLNATLEOLKMAKTNLESAINQANTDKTTDFNEHPNL-----VEAYKALKT	108		
Db	290	LQNLAAQSLEELDKATTN-----EQATQVKNOFLENAQKLKEIQPLIKETNVKLYKAMSE	344		
QY	109	TEQRAFNLEGLSTAYNQIRNNLVLYNKASSLITKTLPLDPLNGGTLTLLDSNEITANKNI	168		
Db	345	SLEQVEKELKHNSA-----NLEDLVAKSKEIVREYEGKLNQSKNLP--ELKQLEEEA	395		
QY	169	NNTLSTINEQ-----KTADALSNSFTKKVI---QNNQESFVGTFTNANVQPSNYSFVAF	220		
Db	396	HSLKQVQVEDFRKKFKTSEQVTPKKRVKRDLAANNENQOKI-----ELTVSPENI-----	445		
QY	221	SADVTVPVNYKYARTVWNGDEPSSRIILANTNSITDYSWIYSLAGTNKYQFSF-----N	275		
Db	446	-----TVYEGEDVKFTVTAKSDSKTLDPSDLL--TKYNPSVSDRISTN	487		
QY	276	YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGVQOQVEFATSTANNTP--TPAV	331		
Db	488	YKNTDNHKKIARITIKLNKLSQVTLKAKDDSGNVVKEFTTIVQKKKEQKVPKTPQE	547		
QY	332	DEIKVAKIV 340			
Db	548	KDSKTEKV 556			
RESULT 8					
RBP2_PLAVB STANDARD; PRT; 1251 AA.					
ID	RBP2_PLAVB				
AC	Q00799;				
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	Reticulocyte binding protein 2 (Fragment).				
GN	RBP2.				
OS	Plasmodium vivax (strain Belem).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=31273;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92315338; PubMed=1617731;				
RA	Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;				
RT	"A reticulocyte-binding protein complex of Plasmodium vivax				
RT	merozoites.";				
RL	Cell 69:1213-1226(1992).				
CC	-1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO				
CC	HUMAN RETICULOCYTE CELLS.				
CC	-1- SUBCELLULAR LOCATION: Membrane-bound (Probable).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M88098; AAA29744.1; -				
KW	Malaria; Receptor; Membrane.				
FT	NON_TER 1 1				
FT	NON_TER 1251 1251				
SO	SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;				
Query Match 6.4%; Score 129; DB 1; Length 1251;					
Best Local Similarity 21.1%; Pred. No. 6;					

FT	DOMAIN	587	631	LAMININ EGF-LIKE 7.
FT	DOMAIN	632	676	LAMININ EGF-LIKE 8.
FT	DOMAIN	677	731	LAMININ EGF-LIKE 9.
FT	DOMAIN	732	784	LAMININ EGF-LIKE 10.
FT	DOMAIN	785	815	LAMININ EGF-LIKE 11 (INCOMPLETE).
FT	DOMAIN	816	1374	DOMAIN IV''.
FT	DOMAIN	1375	1420	LAMININ EGF-LIKE 12.
FT	DOMAIN	1421	1465	LAMININ EGF-LIKE 13.
FT	DOMAIN	1466	1513	LAMININ EGF-LIKE 14.
FT	DOMAIN	1514	1564	LAMININ EGF-LIKE 15.
FT	DOMAIN	1565	1574	LAMININ EGF-LIKE 16 (N-TERMINAL).
FT	DOMAIN	1575	1775	LAMININ DOMAIN IV (DOMAIN IV).
FT	DOMAIN	1776	1808	LAMININ EGF-LIKE 16 (C-TERMINAL).
FT	DOMAIN	1809	1858	LAMININ EGF-LIKE 17.
FT	DOMAIN	1859	1916	LAMININ EGF-LIKE 18.
FT	DOMAIN	1917	1969	LAMININ EGF-LIKE 19.
FT	DOMAIN	1970	2016	LAMININ EGF-LIKE 20.
FT	DOMAIN	2017	2063	LAMININ EGF-LIKE 21.
FT	DOMAIN	2064	2111	LAMININ EGF-LIKE 22.
FT	DOMAIN	2112	2671	DOMAIN II AND I.
FT	DOMAIN	2672	2868	LAMININ G-LIKE 1.
FT	DOMAIN	2876	3048	LAMININ G-LIKE 2.
FT	DOMAIN	3055	3223	LAMININ G-LIKE 3.
FT	DOMAIN	3349	3528	LAMININ G-LIKE 4.
FT	DOMAIN	3534	3709	LAMININ G-LIKE 5.
FT	DOMAIN	2178	2249	COILED COIL (POTENTIAL).
FT	DOMAIN	2301	2321	COILED COIL (POTENTIAL).
FT	DOMAIN	2376	2450	COILED COIL (POTENTIAL).
FT	DOMAIN	2541	2676	COILED COIL (POTENTIAL).
FT	DOMAIN	3270	3296	POLY-THR.
FT	DISULFID	273	282	BY SIMILARITY.
FT	DISULFID	275	296	BY SIMILARITY.
FT	DISULFID	298	307	BY SIMILARITY.
FT	DISULFID	310	330	BY SIMILARITY.
FT	DISULFID	333	342	BY SIMILARITY.
FT	DISULFID	335	367	BY SIMILARITY.
FT	DISULFID	370	379	BY SIMILARITY.
FT	DISULFID	382	400	BY SIMILARITY.
FT	DISULFID	403	414	BY SIMILARITY.
FT	DISULFID	405	421	BY SIMILARITY.
FT	DISULFID	423	432	BY SIMILARITY.
FT	DISULFID	435	445	BY SIMILARITY.
FT	DISULFID	448	460	BY SIMILARITY.
FT	DISULFID	450	468	BY SIMILARITY.
FT	DISULFID	470	479	BY SIMILARITY.
FT	DISULFID	482	492	BY SIMILARITY.
FT	DISULFID	495	507	BY SIMILARITY.
FT	DISULFID	497	514	BY SIMILARITY.
FT	DISULFID	516	525	BY SIMILARITY.
FT	DISULFID	528	538	BY SIMILARITY.
FT	DISULFID	541	553	BY SIMILARITY.
FT	DISULFID	543	560	BY SIMILARITY.
FT	DISULFID	562	571	BY SIMILARITY.
FT	DISULFID	574	584	BY SIMILARITY.
FT	DISULFID	587	599	BY SIMILARITY.
FT	DISULFID	589	605	BY SIMILARITY.
FT	DISULFID	607	616	BY SIMILARITY.
FT	DISULFID	619	629	BY SIMILARITY.
FT	DISULFID	632	644	BY SIMILARITY.
FT	DISULFID	634	650	BY SIMILARITY.
FT	DISULFID	652	661	BY SIMILARITY.
FT	DISULFID	664	674	BY SIMILARITY.
FT	DISULFID	677	691	BY SIMILARITY.
FT	DISULFID	679	700	BY SIMILARITY.
FT	DISULFID	702	711	BY SIMILARITY.
FT	DISULFID	714	729	BY SIMILARITY.
FT	DISULFID	732	746	BY SIMILARITY.
FT	DISULFID	734	753	BY SIMILARITY.
FT	DISULFID	755	764	BY SIMILARITY.
FT	DISULFID	767	782	BY SIMILARITY.
FT	DISULFID	1375	1387	BY SIMILARITY.
FT	DISULFID	1377	1394	BY SIMILARITY.
FT	DISULFID	1396	1405	BY SIMILARITY.

QY 278 -----PSTGYLYFPYKLVRAADANNVGLQKLNNGVQ 310
Db 678 KGTITDELNDYWKQPOQVRSTNESLFTGTGTPMSSYK-----ANPVISPY--SSSHLR 729
QY 311 QVEFATYSTANTTANPPAVDEIKVAKIVLSGLRF-----GONTIELS-- 354
Db 730 QTSNATNTNPMHPQSLAATLNDPSLQSFVRSGSFYSAPQANSQNNINGETENISPR 789
QY 355 -----VPTGEGNMKNKVPAMIGNIYLSSNENNADKI 384
Db 790 ISSDFNLLVPNLSPRLSNDVPIVPGNNITLTTPSHSNILTMNHPTADNI 838
RESULT 13
PST1_YEAST
ID PST1_YEAST STANDARD; PRT; 444 AA.
AC Q12355;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Protoplast secreted protein 1 precursor.
GN PST1 OR YDR055W OR D4214 OR YD9609.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96381250; PubMed=8789263;
RA Brandt P., Ramlow S., Otto B., Bloesker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
of Saccharomycetes cerevisiae chromosome IV";
RL Yeast 12:85-90(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION.
RX MEDLINE=99251092; PubMed=10234784;
RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
RT "Two-dimensional analysis of proteins secreted by Saccharomycetes
cerevisiae regenerating protoplasts: a novel approach to study the
cell wall";
RL Yeast 15:459-472(1999).
RN [4]
RP GPI-ANCHOR.
RX MEDLINE=20469049; PubMed=11016834;
RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
RT "Up-regulation of genes encoding glycosylphosphatidylinositol
(GPI)-attached proteins in response to cell wall damage caused by
disruption of FKS1 in Saccharomycetes cerevisiae";
RL Mol. Gen. Genet. 264:64-74(2000).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
SECRETED BY REGENERATING PROTOPLASTS.
CC -1- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X84162; CAA58971.1; -
DR EMBL; 274351; CAA98873.1; -
DR EMBL; 249209; CAA89084.1; -
DR SGD; S0002462; PST1.
DR COMPIUVEAST-2PAGE; Q12355; -
KW Glycoprotein; Membrane; GPI-anchor; signal.
FT SIGNAL

OX NCBI_TaxID=4932;
RN SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
[3]
RN FIBRONECTIN TYPE III DOMAIN.
RX MEDLINE=97148176; PubMed=8994808;
RA Bateman A., Chothia C.;
RT "Fibronectin type III domains in yeast detected by a hidden Markov
model";
RL Curr. Biol. 6:1544-1546(1996).
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U18779; AAB64999.1; -
DR PIR; S30834; S30834.
DR SGD; S0000769; YEL043W.
DR InterPro; IPR003961; FN_III.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR Hypothetical protein.
KW DOMAIN 35 125 FIBRONECTIN TYPE-III.
FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;

Query Match 6.3%; Score 127; DB 1; Length 956;
Best Local Similarity 19.7%; Pred. No. 5.5;
Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;

QY 5 ITKDNPNNGQTGLEARMELTDL-----INA-KAMTLASLDYAKTEASLSSAYSEA 57
Db 337 LNESNASVANINKEIESLQNEISKMEESNKRNLNASKSLITSIVNANVENDKPIASGEL 396
QY 58 ETYNNNLN-ATLE-----OLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALK 107
Db 397 SAVLKUNDFTLEKNGFSLNAGEEFLSKLNADSLIKMIKQELSIDQE---LEANKLQR 453
QY 108 TTLQRATNLEGLSTAYNGIRNLDVYNKASLITKTL-DPL--NGGTLDSNEITTA 164
Db 454 SNLLKKISALE-----NQF--NENSLNRR--NLKTKLVQPKYKNGDSLAATNSNNSA 502
QY 165 NKN-----INNTLSTINEQKTNADAL----- 185
Db 503 EKNRSSGSIQLPLSNNSMRTGSDILSNKNSIN--NSNADSAAPPLKLNPNVYSFSNRP 560
QY 186 ---SNSFTKKVIONNE--QSFVGTFTNAN---VQPSYSEFA-----FSADVTPVNYK 230
Db 561 IQPSSLLSQTQTDNRMSLNSHISNNENKQPPSSYHALPTATATATATATATIN-G 619
QY 231 YARTVWNG---DEPSSGRILA-----NTNSITDVSWIYSLACTNKKYQFSSNYG 277
Db 620 HSRSLNLTWTAQFAQPSHQVQTELDQAFEDYDANHL--ISGLQNMYYDETDPDINSNYS 677

FT	CHAIN	20	?	PROTOPLAST SECRETED PROTEIN 1. REMOVED IN MATURE FORM (POTENTIAL). SER-RICH.
FT	PROPEP	356	444	
FT	DOMAIN	416		
FT	CARBOHYD	57	57	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	76	76	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	83	83	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	210	210	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	228	228	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	235	235	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	263	263	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	280	280	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	305	305	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	444 AA;	45776 MW;	230860CACAS92IA4 CRC64;
Query Match 6.3%; Score 126.5; DB 1; Length 444;				
Best Local Similarity 18.1%; Pred. No. 2.3;				
Matches 73; Conservative 79; Mismatches 146; Indels 105; Gaps 15;				
QY	17	TOLEAARME-LTDLINAKAMTSLASLODYAKIEASLSAYSEAEVYNNNNLNATLBOLEKMAK	75	
DB	88	TSPAADSLESITDSLNLQSLTILT-----SASFGSLQSVDSIKLITLPAISSFT	136	
QY	76	TNLESATNQANTDKYTFDNEHPNLVEAYKALKTTLEORATNLEGLSS--TAYNOIRNNLV	133	
DB	137	SNIKASNNIYISDTSL-----QSVDFGFSALKYVNVFVNNNNKKLTSIKSPVETVSDSLQ	190	
QY	134	DLYNKASSLTKTLDPL-----NGGTLLOSNETTANKN-----INTLSTINEOKTNA	182	
DB	191	FSFNGNFKI--TFDDLWANNISLTDVHVSFANLQKINSGLGFINNLSISLFTKLNT	248	
QY	183	-----DALSNSFKKK-----VIQNN-----EQSPVGTFTN	207	
DB	249	IGQTFSVISNDYLNLSFNLSLTIGGALVWANNNTGLOKIGGLDMLTTTGGTLEVVGNFTS	308	
QY	208	ANYQPSNYSVAFSADYTPVNYKARYTVWNGDEPSSRIILANTNSITDVSIYSLAGTNT	267	
DB	309	LNL--DSLKSVKGADYESSNSFNAL-----KALQKGGIKGESVFCNGASST	358	
QY	268	KYQPSFNYPGSTGLYFPYKLVKAADANNVGLQKLNNGVQVQFATSTSANNTTAMP	327	
DB	359	SVKLSTSKSQSS-----QTTAKVSKSSKAEKKTSIGDIKAAASASSVSSSGASS--	410	
QY	328	TPAVDEIKAVKIVLUSGRFGONTIELSVPTGEGNNKVPAMIG	370	
DB	411	-----SSSKSGKGNAAIMAPIGQ-----TTPLVG	434	
RESULT 14				
MSPI_PLAFK STANDARD; PRT; 1630 AA.				
ID	MSPI_PLAFK	AC	P04932;	
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Merozoite surface protein 1 precursor (Merozoite surface antigens)			
DE	(PMMA) (p190).			
GN	MSP-1.			
OS	Plasmodium falciparum (isolate KI / Thailand).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5839;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86136024; PubMed=3004972;			
RA	Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certu U.,			
RA	Stunnenberg H., Bujaard H.			
RT	"Polymorphism of the precursor for the major surface antigens of			

QY 328 TPAVDEIKVAKIVLGLRFGONTIELSVPTGEGNMKVAPMIGNIYLSNNENADK 383
 Db 549 IVVEKELKYKKNLSKI-----ENEIETLVE-----NIKDEQLFEKKTIDENKPE 597

RESULT 15
 MSP1_PLAFW
 ID MSP1_PLAFW STANDARD; PRT; 1639 AA.
 AC P04933;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMWSA) (P195).
 GN MSP-1.
 OS Plasmodium falciparum (isolate Wellcome).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=5848;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=86014355; PubMed=2995820;
 RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
 RA Nicholls S.C., Hillman I., Davey L.S., Tizard M.L.V., Schwarz R.T.,
 RA Freeman R.R.;
 RT *Primary structure of the precursor to the three major surface
 RT antigens of plasmodium falciparum merozoites.*;
 RN Nature 317:270-273(1985).
 RL [2]
 RP REVISIONS.
 RA Holder A.A.;
 RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
 CC -|- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -|- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X02919; CAA26676.1; -.
 DR PIR; A24594; A24594.
 DR InterPro; IPR000561; EGF-like.
 DR Pfam; PF00008; EGF; 1.
 KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
 FT Transmembrane; GPI-anchor.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match 6.3%; Score 126.5; DB 1; Length 1639;
 Best Local Similarity 20.7%; Pred. No. 11;
 Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

QY 37 LASLDYAK-----IEASLSAYSEATVNNNLNATLEQLKMAKTNLESAINOANTDKT 91
 Db 235 VGMEDYIKKNKKTIEINELIEESKKTIDKNKNATKEEK-----KKLYOQAYDLSI 287

QY 92 FDNEHPLVAYKALKTTLEQATNLEGLSSTAYNIRNNLVLYNKASLIKTILDPN 151
 Db 288 YNKQ---LEEAHN-LISVLEKRIIDLK-----NENIKELDKINEI---KNPPAN 332

QY 152 GG---TLDSN-EITTANKNINNTLSTINEQKTNADAL-----SNSFIK 191
 Db 333 SGTPNTLLDKNKKIEHEKEIEIAKTI---AFNIDSLFDPLELEYLRKKNKIDIS 389

QY 192 KVIONNESQSVGFTTNANVQPSNYSFV-----AFSADVTVPVNYKYARTVWNGDEPS 243
 Db 390 AKVETKSTEPNEYPNGVTYPLSYNDINNALNELNFGDLINPFDYT-----KEPS 440

QY 244 SRLAN-----TNSITDVSWIYSLAGTNTKYQFS 272
 Db 441 KNIYTDNERKKFNEIEKEKIEKKKIESDKSYEDRSKSLNDITKEYKL-LNEYDYSK 499

QY 273 FSNYGPSTGY-----LYFPYKLVKAADANNVGLQYKLNNGNVQOVFATSTSANNTANP 327
 Db 500 FNNNIDLTNFEKMMGKRYSYKVEKLTH-HNTFASYENSKHLEKLTALKYMEDYSLRN- 557

QY 328 TPAVDEIKVAKIVLGLRFGONTIELSVPTGEGNMKVAPMIGNIYLSNNENADK 383
 Db 558 IVVEKELKYKKNLSKI-----ENEIETLVE-----NIKDEQLFEKKTIDENKPE 606

Search completed: June 12, 2002, 10:51:09
 Job time: 207 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	831.5	41.5	702	2	S48754	major surface prot
2	754	37.6	702	2	S48753	major surface prot
3	744	37.1	650	2	S48751	major surface prot
4	743	37.1	649	2	S48752	major surface prot
5	713	35.6	647	2	A49218	hemagglutinin homo
6	575.5	28.7	386	2	S48755	major surface prot
7	189.5	9.5	320	2	S51560	major surface prot
8	175	8.7	6713	2	B89921	hypothetical prote
9	173	8.6	1302	1	JG5009	surface-located me
10	165.5	8.3	661	2	AC2422	hypothetical prote
11	164.5	8.2	2481	2	D90011	FmTB protein [mpo
12	163.5	8.2	3890	2	C89921	hypothetical prote
13	159	7.9	933	2	S41539	fibrinogen-binding
14	156	7.8	807	2	B71605	hypothetical prote
15	155	7.7	1072	2	A86827	hypothetical prote
16	155	7.7	4688	2	F83288	hypothetical prote
17	154.5	7.7	682	2	S44131	subtilisin-like pr
18	154.5	7.7	1365	2	T30822	lmp1 protein - Myc
19	151.5	7.6	1237	2	D71850	probable outer mem
20	151	7.5	624	2	PC6003	surface membrane p
21	150.5	7.5	1051	2	T18351	lmp1 protein - Myc
22	145	7.2	2399	2	H71879	toxin-like outer m
23	144.5	7.2	1487	2	ACG560	hypothetical prote
24	144	7.2	1107	2	AC0976	probable autotrans
25	143.5	7.2	751	2	T40462	ser-lys rich hypot
26	142	7.1	135	2	B49218	hemagglutinin homo
27	142	7.1	2860	2	E85822	probable invasiv z
28	141.5	7.1	1645	2	F96907	phage-related prot
29	140.5	7.0	3194	2	D71917	toxin-like outer m

A:Cross-references: GB:D90714; EMBL:L28423; NID:g1905869; PIDN:AA650152.1; PID:g19058
R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whithear, K.G.
Infect. Immun. 60: 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepti
A:Reference number: A44793; MUID:92363591
A:Contents: S6
A:Accession: A44793
A:Status: preliminary
A:Molecule type: protein
A:Residues: 26-42 <MA>
A>Note: sequence extracted from NCBI backbone (NCBIP:111017)
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

RESULT 2
S48753
major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface P
A:Reference number: S48751; MUID:95010739
A:Accession: S48753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-702 <MA>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.6%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 6.2e-34;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;
QY 11 NPNNGT---OLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNNLNA 66
Db 69 NPNNGTTPPEQLAARLTLDLGTENTVALYADYAKIQSTLSTAYMTAKTASENTSA 128
QY 67 TLEQLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLSSAYN 126
Db 129 TLENLSASTTLOAAIDRAANDKRVDSVNPQPLVAAAYNNLTKLTKSTTSLEGLSENKYG 188
QY 127 QTRNLVLYNKASSLTKPLDPLNG--GTLLDSNEITTANKNNINTLS--TINEQKINA 182
Db 189 GIKNHLKSLFDGSAITAKTLDPTGERTPLEKYNE--ANNKIMALSPESLKKWKGNA 245
QY 183 DALSNSFIKKVTONNEQSPVGTFTNANQPSNYSFVAFSADVTP-----VNYKARRIVW 237
Db 246 DKF-NEFEKNPLSKELKSTDAHNQEPANWSPAAYSVDLTSNSQNLPMNFAQRKVV 304
QY 238 NGD--PPSSRLANTN-STIDVSWIYSLAGTNTKYQFSNYPSTGYLYPYKLVKAAAD 294
Db 305 TSENOQPGKALTSPVSDVSWIYSLAGTNTKYQFSNYPSTGYLYPYKLVKAAAD 364
QY 295 ANNVGLQYKINGNVQOQVEF-----ATST-----SAN 321
Db 365 SSSVALQYSLNKTSSKLINFKPAETVSTNTDQSENEVATTSTEARSSYKVLVADAAATS 424
QY 322 NTPANTPPADEVKIVLGLRFGQNTIELSVPTGEGNMKNKVPIMGNIVLSSNENNA 381
Db 425 NNEHNTPTVSDINIAKVTLSGLTGFENTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479
QY 382 DKI 384
Db 480 VKI 482

RESULT 3
S48751
major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum
N:Alternate names: major hemagglutinin pmGA
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface P
A:Reference number: S48751; MUID:95010739
A:Accession: S48751
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-650 <MA>

Query Match 37.1%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 2e-33;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;
QY 11 NPNNGQ-----TOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNNN 63
Db 51 NPGDGGOMNAAQAELAAARMGLTTTFVDSKAKNGLGYDYKKTONTLTAKYDAAKTVLDIN 110
QY 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLSST 123
Db 111 SSSITQNLNEAKTRLETAFTAATSAQTQDEQHAELVYKVKELTKLTLSNETATLAPYADA 170
QY 124 AYQIRNRLVDLYNKASSLTKPLDPLNGTLLDSNEITTANKNNINTL--STINEQKIN 181
Db 171 QYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTAGAVTMAATKIYEAIKDEVLPNPKEN 229
QY 182 ADALNSFIKKVTONNEQSPVGTFTNANQPSNYSFVAFSADVTPV-----NYKYAR 233
Db 230 ATKLAISFVQVIVKEKITGVEEAHN-KAOPANYSFVGSVDITGTANGOTSIPNNVYAO 288
QY 234 RTVM-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGYLY 284
Db 289 RTITNGDEP-RSVSNTPVDGQTMQAQPLSNVSWIYSLAGTNTKYQFSNYPSTGYLY 346
QY 285 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
Db 347 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
QY 344 LRFGQNTIELSVPTGEGNMKNKVPIMGNIVLSSNENNAKDI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

RESULT 4
S48752
major surface protein (clone pmGAL.2) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S48752
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-649 <MA>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 2.2e-33;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;
QY 11 NPNNGQ-----TOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNNN 63
Db 51 NPGDGGOMNAAQAELAAARMGLTTTFVDSKAKNGLGYDYKKTONTLTAKYDAAKTVLDIN 110
QY 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLSST 123
Db 111 SSSITQNLNEAKTRLETAFTAATSAQTQDEQHAELVYKVKELTKLTLSNETATLAPYADA 170
QY 124 AYQIRNRLVDLYNKASSLTKPLDPLNGTLLDSNEITTANKNNINTL--STINEQKIN 181
Db 171 QYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTAGAVTMAATKIYEAIKDEVLPNPKEN 229
QY 182 ADALNSFIKKVTONNEQSPVGTFTNANQPSNYSFVAFSADVTPV-----NYKYAR 233
Db 230 ATKLAISFVQVIVKEKITGVEEAHN-KAOPANYSFVGSVDITGTANGOTSIPNNVYAO 288
QY 234 RTVM-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGYLY 284
Db 289 RTITNGDEP-RSVSNTPVDGQTMQAQPLSNVSWIYSLAGTNTKYQFSNYPSTGYLY 346
QY 285 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
Db 347 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
QY 344 LRFGQNTIELSVPTGEGNMKNKVPIMGNIVLSSNENNAKDI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

RESULT 5
S48751
major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum
N:Alternate names: major hemagglutinin pmGA
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S48751
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-650 <MA>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 2.2e-33;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;
QY 11 NPNNGQ-----TOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNNN 63
Db 51 NPGDGGOMNAAQAELAAARMGLTTTFVDSKAKNGLGYDYKKTONTLTAKYDAAKTVLDIN 110
QY 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLSST 123
Db 111 SSSITQNLNEAKTRLETAFTAATSAQTQDEQHAELVYKVKELTKLTLSNETATLAPYADA 170
QY 124 AYQIRNRLVDLYNKASSLTKPLDPLNGTLLDSNEITTANKNNINTL--STINEQKIN 181
Db 171 QYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTAGAVTMAATKIYEAIKDEVLPNPKEN 229
QY 182 ADALNSFIKKVTONNEQSPVGTFTNANQPSNYSFVAFSADVTPV-----NYKYAR 233
Db 230 ATKLAISFVQVIVKEKITGVEEAHN-KAOPANYSFVGSVDITGTANGOTSIPNNVYAO 288
QY 234 RTVM-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGYLY 284
Db 289 RTITNGDEP-RSVSNTPVDGQTMQAQPLSNVSWIYSLAGTNTKYQFSNYPSTGYLY 346
QY 285 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
Db 347 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
QY 344 LRFGQNTIELSVPTGEGNMKNKVPIMGNIVLSSNENNAKDI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

RESULT 6
S48751
major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum
N:Alternate names: major hemagglutinin pmGA
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S48751
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-650 <MA>

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 2.2e-33;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;
QY 11 NPNNGQ-----TOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNNN 63
Db 51 NPGDGGOMNAAQAELAAARMGLTTTFVDSKAKNGLGYDYKKTONTLTAKYDAAKTVLDIN 110
QY 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLSST 123
Db 111 SSSITQNLNEAKTRLETAFTAATSAQTQDEQHAELVYKVKELTKLTLSNETATLAPYADA 170
QY 124 AYQIRNRLVDLYNKASSLTKPLDPLNGTLLDSNEITTANKNNINTL--STINEQKIN 181
Db 171 QYAGIKMHLGLYDAGKAITTKTLEPEVGGP-LTAGAVTMAATKIYEAIKDEVLPNPKEN 229
QY 182 ADALNSFIKKVTONNEQSPVGTFTNANQPSNYSFVAFSADVTPV-----NYKYAR 233
Db 230 ATKLAISFVQVIVKEKITGVEEAHN-KAOPANYSFVGSVDITGTANGOTSIPNNVYAO 288
QY 234 RTVM-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGYLY 284
Db 289 RTITNGDEP-RSVSNTPVDGQTMQAQPLSNVSWIYSLAGTNTKYQFSNYPSTGYLY 346
QY 285 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
Db 347 PPKLVKAAADANNVGLQYKINGNVQOQVEFATSTANNTTAN-PTPAVDKIVKIVLVS 343
QY 344 LRFGQNTIELSVPTGEGNMKNKVPIMGNIVLSSNENNAKDI 384
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

```

RESULT 7
S51560
major surface protein (clone pMGAl.6) - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S51560; S48757
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surfac
A:Reference number: S48751; MUID:95010739
A:Accession: S51560
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA

```

Fri Oct 10 11:43:24 2003

Qy 364 -----KVAPMIGNIYLSNEN 379
Db 886 LOGAINDKDATLRNQNYLDADES 908

RESULT 9
JC6009
surface-located membrane protein lmp3 precursor - Mycoplasma hominis
C:Species: Mycoplasma hominis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: JC6009
R:Adelofed, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.
J. Bacteriol. 178, 2775-2784, 1996
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene sys
A:Reference number: JC6009; MUID:96213016
A:Accession: JC6009
A:Molecule type: DNA
A:Residues: 1-1302 <LAD>
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336
C:Genetics:
A:Gene: lmp3

A:Genetic code: SGC3
C:Superfamily: surface-located membrane protein lmp3; tetratrico peptide repeat homolo
C:Keywords: duplication; membrane protein
F:1-24/Domain: signal sequence status predicted <SIG>
F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
F:957-992/Domain: tetratrico peptide repeat homology <Tn1>
F:993-1026/Domain: tetratrico peptide repeat homology <Tn2>
F:1089-1120/Domain: tetratrico peptide repeat homology <Tn3>
F:1154-1190/Domain: tetratrico peptide repeat homology <Tn4>

Query Match 8.6%; Score 173; DB 1; Length 1302;
Best Local Similarity 22.8%; Pred. No. 0.08;
Matches 94; Conservative 67; Mismatches 163; Indels 88; Gaps 18;
Qy 5 ITTKDAMPNGQ-----TQLEARMELTDLINAKAMTLASLDYAKIEASLSAYSAETV 60
Db 864 ITTKLETFNKDKVKFKELEQTRKIDDEFIN-----TNKTPDYSTLISELTSKRDSKNSI 919

Qy 61 NNUNATLEQLKMAKNLESAINQANTDKTFDN-----EHPN-----LVEAYKAL 106
Db 920 TNSNKS--DIETANTEKQALAKANTDKQADNLAARSTREQLNKSISANTILLAKLTKR 977
Qy 107 KTTLEQATNLEGLSSAYNQI--RNNDVLYNKASSL-----ITKPLDPLNGTLLDSN 159
Db 978 DNTIOQAKTELEKEVQKQANQAVASNTASMQSAKSLDAKVTETTKLETFNKDKVKER 1037
Qy 160 EITFANKNNITLSTINEQKTNADALSNSFKKVIQNEQSFVGTFTTNANVQPSNYSEVA 219
Db 1038 ELEOTRKIDDEFINT--NKTNPYSTLISELTSK--RDSKNSITNSNKSIDETANTEL-- 1092
Qy 220 FSADVPVNYKYARTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFNYGPS 279
Db 1093 -----KQALAKAN--TDKAQADNLA--RSTKEQLNKSISAN 1125
Qy 280 TCYLYFPYKLVKAADANNVGLYKLNNGVQVQVEFATSTANNTTANPTPAVDEI--KVA 337
Db 1126 T-----LLAKLTDKNTIQAKTE--LEKEVQKQANQAVASNTASMQSAKSLDTKVT 1176
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNKNKVPAMIGNI--YLSNENNAADKIPCY 387
Db 1177 EITKLETFNKDK-----EAKFNELKTRGQIEFINTNKN-----PNY 1216

RESULT 10
AG2422
hypothetical protein all14935 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AG2422
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

A:Residues: 1-320 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62415.1; PID:g535688
C:Genetics:
A:Genetic code: SGC3

Query Match 9.5%; Score 189.5; DB 2; Length 320;
Best Local Similarity 48.9%; Pred. No. 0.0016;
Matches 46; Conservative 11; Mismatches 26; Indels 11; Gaps 3;
Qy 292 AADANNVGLQYKLVNGVQVQVEFATSTANNTTANPTPAYDEIKVAKIVLSGLRFGQNTI 351
Db 41 SADS-----PTQENSQSNQAPEASA-----MNETPTVDGINVAKVITLTLKFGSNTI 90

Qy 352 ELSVP--TQEGNNKVPAMIGNIYLSNENNAADKI 384
Db 91 ELSVPTTDEGTSKVPAMIGNIYTSNDQNKRI 124

RESULT 8
B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: B89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Query Match 8.7%; Score 175; DB 2; Length 6713;
Best Local Similarity 22.6%; Pred. No. 0.55;
Matches 100; Conservative 66; Mismatches 179; Indels 98; Gaps 17;
Qy 9 DANNPQGTQLEARMELTDLINAKAMTLASLDYAKIEASLSAYSAET-----VNNN 63
Db 492 EAVSNTG--TQNTAMANLQNGINDKANTLAS--ENYHDADSKKTATYQAVTNAENILAKN 549

Qy 64 LNATLEQLKMAKNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEQATNLEGLST 123
Db 550 SGNLND-----KAAVENALSQVNTAKGALNGNH--NLEQAKSNANTTI-----NGLQHLTPA 599
Qy 124 AYQIRNVLVDLYNKAS--SLITKTLPLNG--GTLDSNEITANKNNITLSTINEQKT 180
Db 600 QKDLKQVQQAQNAVAGVDTKSSANTLNGAMGTLENSIQDNTATNGQNYLDATERNKT 659
Qy 181 -----NADALSFTKKVIQNEQSFVGTFTTNANVQPSNYSEVA 221
Db 660 NYNNAVDSANGVINATSNPNNDANAIQIATQVTSKALDGHNLTOAKQT-----ATN 714
Qy 222 ADVTPVNYKYARTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFNYGPS 279
Db 715 AIDGATNLNKAQDALKAAQVTSQORVANVTSIQOOTANELNTAMGQLOHGIDDENATKQY 774
Qy 270 QFSFNSGPTGYLYFPYKLVKAADA-----NNVGLQYKLN-----NGNVQ 310
Db 775 KYRDEQSKKTAY----DQAVAAKAILNKQGTGNSDKAAVDRALQOQVTSKDALNGDAK 830
Qy 311 QVEFATSTSAN-----NTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNKN-- 363
Db 831 LAEAKAARQNLGTLNHTNAQRTALEGOINQATV-----DGVNTVTKTNTANTLDGAMNS 885

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-661 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076634.1; PID:g17134073; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all14935

Query Match 8.3%; Score 165.5; DB 2; Length 661;
Best Local Similarity 22.7%; Pred. No. 0.084;
Matches 88; Conservative 40; Mismatches 188; Indels 71; Gaps 11;
QY 19 LEARMELTDLI-----NAKAMTLASLDYAKTEASLSNAYSEATVNNLNATLEQ 70
DB 64 LKKAEDDLGIIRSTNRKSDANSRLNRAEQDFTQADGFOAQATATNTATNNRNAQNA 123
QY 71 LKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSSYAYNOIRN 130
DB 124 LNTATNNRNAQNALNTATNNRDN-----AQNALNTATNNR--NNAQNALNTATNN--RN 174
QY 131 NLVDLYNKASSLTTKTDLPLNGGTLDSNEITPANKNINNTISTINEQKTNADALSFTI 190
DB 175 NAQNALNTATNNRNAQNAL-----NTATNNRNAQNALNTATNNRNAQAEVDAT 226
QY 191 KVIQ---NNEQSFVCTFTFNANVQPSNYFVAFSADVTVPYKYARRTWNQDEPSSRIL 247
DB 227 RNLQAARRGNSQRAIQAINALNQANRLNTANTALNTATNNFNNTANTAFN-----T 278
QY 248 ANTNSITDVSIVSYSLAGTNTKYQSFSPNYGPGTGYLYFPYKLVKAADANVGYQKLNNG 307
DB 279 ANNFST-----ANTAFNTATNNFSTA-----NTALNTATNNF 311
QY 308 NVOQVEFATSTANNTANPTPAVDKIKVILSGLRFGQNTIELSVPTGEGNMKNKVP 367
DB 312 NTAIELDQANRLTARN-----DFNTANSNFSRTGNELNTATNNFNNTANTNTAIT 365
QY 368 MIGNIYLSNENNADKIPGVRPGTFL 394
DB 366 NFNN--ASSRNTAEQARNQVRETRL 390

RESULT 11
D90011
FmC protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D90011
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2481 <KUR>
A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BA043253.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fntB(mrp)

Query Match 8.2%; Score 164.5; DB 2; Length 2481;
Best Local Similarity 23.1%; Pred. No. 0.55;
Matches 98; Conservative 56; Mismatches 156; Indels 115; Gaps 20;

QY 9 DANPNNGQTGLE-----AARMELTDLINAKAMTLASLDYAKTEASLSAYSEAE 58
DB 1186 DOAKTTGETSIDQVTPVNVKATARNETAILNNKLEIQATPDATDEKQAADA--EAN 1243
QY 59 TVNNNLN-----ATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTL-- 110
DB 1244 TENGKANQALSAATTNAQVDEAKANAIAIN-AVTPKVVKKQAQKDEIDOLQATQNTVIN 1302
QY 111 -EQRATNLEGLS-----STAYNQIRNNL-----VDLYNKASSLTITKTLPLNGGTL 156
DB 1303 NDONATNEEKEAIIQOLATAVTADKNNITAAATDDNGVDYAKAGKNSIQSTQ--ATAV 1359
QY 157 DSN-----EITANKNINNTLSTNEOKTNADALSNSFIKKVIONNQSFVGTFTNAN 209
DB 1360 KSNKNEVDQAVTQNAIDNTTGTATTEKNAKDL-----VLKAKEKAYQDIL---N 1409
QY 210 VQPSNYFVAFSADVTVPYKYARRTWNQDEPSSRLANTNSITDVSIVSYSLA---GT 265
DB 1410 AQTN-----DVTQI-----KQAVADIQGITADTTIKVAKDELAT 1446
QY 266 NTKYQSFSPNYGPGTGYLYFPYKLVKAADAN-----NVGLQYKLNNGVQVFEFATST 318
DB 1447 KANEQKAL-----IAQTADATTEEKQANQVDAQLTQGN-QNIENAQSI 1490
QY 319 SANNTA-NPTPAVDKIKVILSGLRFGQNTIELSVPTGEGNMKNKVPIMIGNIYLSN 377
DB 1491 DVNTAKDNAIQADIPQASTDVTKNAR-----AELLTEM-----QNKITEILNNNETTNE 1541
QY 378 ENNAD 382
DB 1542 EKGND 1546

RESULT 12
C89921
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <KUR>
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BA042528.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhB

Query Match 8.2%; Score 163.5; DB 2; Length 3890;
Best Local Similarity 24.4%; Pred. No. 1.1;
Matches 94; Conservative 54; Mismatches 153; Indels 85; Gaps 20;
QY 17 TQLEAARMELTDLINAKAMTLASLDYAKTEASLSAYSEA-ETVNNLNATLEQLKMAK 75
DB 3528 TDLTAMGNLQAGINDEQTTLS-QNYQDATPSKKTATYNAVQAQKDLNKSNGQNK-TK 3585
QY 76 TLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSTAYNQIRNNLVDL 135
DB 3586 DQTEAMNQVNSAKNLDG-----TRLLDQAKQAKQQLNNHTLTQAQNTLNQ----- 3636
QY 136 YNKASSLTITKTLPLNGGTLDSNEITANKN-INNTLSTINEQKTNADA--LSNSFI-- 190
DB 3637 -----INSGLTVAGVHTVQSNANTLDQAMNTLROSIAANNNDATKASEDYDA 3682
QY 191 ---KKVIONNEQSFVGTFTFNANVQPS-NYSFVAFSADVTVPYKYARRTWNQDE----- 241

RESULT 14
B71605

RESULT 15

A86827
hypothetical protein yafg [imported] - Lactococcus lactis subsp. lactis (strain IL140)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 2301 #text_change 03-Aug-2001
C:Accession: A86827
R:Botolin, A.; Wincker, P.; Mauer, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86825; MUID:21235186; PMID:11337471
A:Accession: A86827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yafg

Query Match 7.7%; Score 155; DB 2; Length 1072;
Best Local Similarity 21.0%; Pred. No. 0.6;
Matches 97; Conservative 76; Mismatches 154; Indels 136; Gaps 20;
9 DANPNNGOTQLEARMELTDLINAKMTLASLDYAKIEASLSAYSE-AETVNNNL--- 64

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:49:08 ; Search time 22.56 Seconds
(without alignments)
426.582 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTTKDAMPNGQTGLE.....SSNNENADKIPYRRGTFPL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	95.3	615	2	US-08-525-742-10
2	1856	92.6	610	2	US-08-525-742-8
3	1638	81.7	368	2	US-08-525-742-4
4	1612	80.4	368	1	US-08-185-851A-4
5	806	40.2	661	2	US-08-525-742-2
6	159	7.9	933	3	US-08-293-728-2
7	159	7.9	933	4	US-09-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1002	4	US-09-268-347-24
10	135.5	6.8	1073	4	US-09-341-782-6
11	134	6.7	1004	4	US-09-268-347-30
12	131.5	6.6	1164	4	US-08-923-992A-2
13	131	6.5	1104	4	US-08-923-992A-4
14	129	6.4	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	122.5	6.1	1536	1	US-08-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	1	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2
					Patent No. 5352450

28	120.5	6.0	746	5	PCT-US95-10509-2	Sequence 2, Appli
29	120.5	6.0	1104	4	US-09-268-347-28	Sequence 28, Appli
30	120.5	6.0	1104	4	US-09-268-347-34	Sequence 34, Appli
31	119.5	6.0	984	1	US-08-242-932-2	Sequence 2, Appli
32	119.5	6.0	984	1	US-08-714-481-2	Sequence 2, Appli
33	119.5	6.0	984	5	PCT-US95-06111-2	Sequence 2, Appli
34	119.5	6.0	2314	4	US-09-268-347-49	Sequence 49, Appli
35	119	5.9	2354	4	US-09-268-347-47	Sequence 47, Appli
36	118.5	5.9	866	2	US-08-483-101-4	Sequence 4, Appli
37	118	5.9	1612	1	US-08-169-927-2	Sequence 2, Appli
38	117.5	5.9	889	4	US-09-336-447A-15	Sequence 15, Appli
39	117.5	5.9	1912	1	US-08-409-995-4	Sequence 4, Appli
40	117.5	5.9	1912	3	US-08-585-467-4	Sequence 4, Appli
41	117	5.8	2353	4	US-09-377-155-33	Sequence 33, Appli
42	117	5.8	2353	4	US-08-913-942-4	Sequence 4, Appli
43	117	5.8	2353	4	US-09-569-974-33	Sequence 33, Appli
44	116	5.8	941	4	US-09-336-447A-9	Sequence 9, Appli
45	116	5.8	1338	2	US-08-728-470-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Sakai, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hiroh

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigemi

; APPLICANT: Takahashi, Kiyoohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; ADDRESSEE: NAUGHTON

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-074139

; FILING DATE: 31-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-245625

; FILING DATE: 30-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/00541

; FILING DATE: 31-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: McLealand, Le-Nhung

; REGISTRATION NUMBER: 31,541

; REFERENCE/DOCKET NUMBER: 950811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-659-2930

```
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-10

Query Match          95.3%; Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%; Pred. No. 3.9e-132;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 61
DB 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTDFDNEHPNLYEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFDNEHPNLYEAYKALKTTLEQRATNLEGLA 146
QY 122 STAYNOIRNNLVLYNKASSLITKTLDPLNGTGLDSDNEITANKNINNTLSTINEQKTN 181
DB 147 STAYNOIRNNLVLYNKASSLITKTLDPLNGTGLDSDNEITANKNINNTLSTINEQKTN 206
QY 182 ADALNSFIKKYIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 241
DB 207 ADALNSFIKKYIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 266
QY 242 PSSRLANTNSTIDVSWIYSLAGTNTKQFSEFSNGPSTGYLYFPYKLVKAADANNVGLQ 301
DB 267 PSSRLANTNSTIDVSWIYSLAGTNTKQFSEFSNGPSTGYLYFPYKLVKAADANNVGLQ 326
QY 302 YKLNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTGEGN 361
DB 327 YKLNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTGEGN 386

QY 362 MNKVAPMIGNIYLSNENNAADKI 384
DB 387 MNKVAPMIGNIYLSNENNAADKI 409

RESULT 2
US-08-525-742-8
; Sequence 8, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-8

Query Match          92.6%; Score 1856; DB 2; Length 610;
Best Local Similarity 95.5%; Pred. No. 3.5e-128;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 61
DB 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTDFDNEHPNLYEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFDNEHPNLYEAYKALKTTLEQRATNLEGLS 146
QY 122 STAYNOIRNNLVLYNKASSLITKTLDPLNGTGLDSDNEITANKNINNTLSTINEQKTN 181
DB 147 STAYNOIRNNLVLYNKASSLITKTLDPLNGTGLDSDNEITANKNINNTLSTINEQKTN 206
QY 182 ADALNSFIKKYIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 241
DB 207 ADALNSFIKKYIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 266
QY 242 PSSRLANTNSTIDVSWIYSLAGTNTKQFSEFSNGPSTGYLYFPYKLVKAADANNVGLQ 301
DB 267 PSSRLANTNSTIDVSWIYSLAGTNTKQFSEFSNGPSTGYLYFPYKLVKAADANNVGLQ 326
QY 302 YKLNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTGEGN 361
DB 327 YKLNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSNENNAADKI 383
DB 387 MNKVAPMIGNIYLSNENNAADKI 408

RESULT 3
US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
```

APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
APPLICANT: Takahashi, Kiyochiko
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
TITLE OF INVENTION: AS USE THEREOF
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-4

Query Match 81.7%; Score 1638; DB 2; Length 368;
Best Local Similarity 95.5%; Pred. No. 1.6e-112;
Matches 322; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 2 CMSITKDPNNGQQLQEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDPNNGQQLQEAARMELTDLINAKARTILASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLA 146
QY 122 STAYNOIRNLVDLYNKASLITKTLDPNGGTLDSNEITTANKNINNTLSTINQKTN 181
DB 147 STAYNOIRNLVDLYNKASLITKTLDPNGGTLDSNEITTANKNINNTLSTINQKTN 206
QY 182 ADALSNFISFKVQNNQEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
DB 207 ADALSNFISFKVQNNQEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 242 PSSRIILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 301
DB 267 PSSRIILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 326

QY 302 YKLNNNGVOQVEPATSTSANNTTANPTPAVDEIKVAK 338
DB 327 YKLNNNGVOQVEPATSTSANNTTANPTQQLMLRLKLLK 363
RESULT 4
US-08-185-851A-4
Sequence 4, Application US/08185851A
Patent No. 5489430
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Okawa, Setsuko
APPLICANT: Fujisawa, Ayumi
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
TITLE OF INVENTION: Well As Vaccines Utilizing the Same
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESSEE: Naughton
STREET: 1725 K Street, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
SOFTWARE: ASCII from Word Perfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,851A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918
TELEPHONE: 202-659-2930
TELEFAX: 202-887-0357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-185-851A-4

Query Match 80.4%; Score 1612; DB 1; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.3e-110;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSITKDPNNGQQLQEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDPNNGQQLQEAARMELTDLINAKARTILASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHNPVLEAYKALKTTLEQRTNLEGLA 146
QY 122 STAYNOIRNLVDLYNKASLITKTLDPNGGTLDSNEITTANKNINNTLSTINQKTN 181
DB 147 STAYNOIRNLVDLYNKASLITKTLDPNGGTLDSNEITTANKNINNTLSTINQKTN 206
QY 182 ADALSNFISFKVQNNQEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
DB 207 ADALSNFISFKVQNNQEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARTVWNGDE 266

QY 242 PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYGSPSTGTYLFFPKYKLKAAADANNVGLQ 301
 Db 267 PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYGSPSTGTYLFFPKYKLKAAADANNVGLQ 326
 QY 302 YKLNGNVQVEFATSTSANNTTANPTTAPVDEIKVAK 338
 Db 327 YKLNGNVQVEFATSTSANNTTANPTTQQLMLRLKLLK 363

RESULT 5
 US-08-525-742-2
 ; Sequence 2, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hirono
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemi
 ; APPLICANT: Takahashi, Kiyoohito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSEE: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742
 ; FILING DATE: 25-SEP-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-074139
 ; FILING DATE: 31-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-245625
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP94/00541
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLeLand, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 950811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-659-2930
 ; TELEFAX: 202-8870357
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-525-742-2

Query Match 40.2%; Score 806; DB 2; Length 661;
 Best Local Similarity 42.8%; Pred. No. 2.9e-51;
 Matches 185; Conservative 62; Mismatches 123; Indels 12;

QY 9 DANPNRGOTQLEAARMELTD-----LINAKAMTILASLODYAKIIBASLSAYSEAEVTN 61
 Db 45 NTPSDGQGMNNAARELADAKAALFTLLINGETANLASVEDYAKIKSELTSAYETAKAVS 104
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNLVEAYKALKTTLEQRTNLEGLS 121
 Db 105 AKTGATLNEYNEAKTTLDAAIKKAASAKNDFDAQHGLSVEAYNNKLTKEETLNLDLSA 164
 QY 122 STAYNOIRNNLVLYNKASSLITKTLPLNGGTLTLLDSNEITTTANKNINNTLSTINEQKTN 181
 Db 165 NENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPVMSVTQANODITNATSLIAWKQN 223
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSYSFVAFSADVTPVNYKYARTVWNGDE 241
 Db 224 ADNLANSFIKQSLVKNNLTRVDVANNOE-OPANYSFVGSVNYDTPNMNFAQKRVWASEN 282
 QY 242 -----PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYG-PSTGYLYFFPKYKLKAAAD 294
 Db 283 TPLATTPAEDATQQAASLTDVSWIYSLNGAEAKYTLTSFRYGAETKTAIYLFYKLVKTS 342
 QY 295 ANNVLQYKLNNGNVQVEF-----ATSTSAN--NTTANP-----TPAVDEIKVAKIV 340
 Db 343 --NVGLQYKLNNGGDTKQINFVQTPASGSDVAANEETMASPAEMOSAPTVDIKIAKVA 400
 QY 341 LSGURFGONTIELSVPTGEGNMKNVAPMIGNIYLS-----NENN 380
 Db 401 LSNLKFSNTIEFSVPTG-----KAAPMIGNMYLTSSSEVNKNKIYDDLFSGNSFNENN 455
 QY 381 -----ADKIPGY 387
 Db 456 PTATVYDILLAGY 467

RESULT 6
 US-08-293-728-2
 ; Sequence 2, Application US/08293728D
 ; Patent No. 6008341
 ; GENERAL INFORMATION:
 ; APPLICANT: Foster, Timothy J.
 ; APPLICANT: McDevitt, Damien L.
 ; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
 ; FILE REFERENCE: 05344.105011
 ; CURRENT APPLICATION NUMBER: US/08/293,728D
 ; CURRENT FILING DATE: 1994-08-22
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 933
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; US-08-293-728-2

Query Match 7.9%; Score 159; DB 3; Length 933;
 Best Local Similarity 22.8%; Pred. No. 0.0011;
 Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY 3 MSITKDKANPNNGQTQL--EAARMELTDLINAKAMTILASLODYAKIIBASLSAYSEAEVT 60
 Db 73 VSDTKTSSNTNNGETSVAAQNPAQOETTSSTNATT-----EETPTVTEATT 120
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNLVEAYKALKTTLEQRTNLEGL 120
 Db 121 TTNQANTPATTSSTNNAELVNO-TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
 QY 121 SST-----AYNQIRNNLVLYNK-----ASSLITKTLPLDPLNGGTL 156
 Db 173 STTQDTSTPATPSNNESAPQSTDSNKKVNOAVNTSAPRMRAFSLAAVAADAPAAAGTDI 232
 QY 157 DSNETTTANKNINNTLSTINEQ----KTN-ADALNSFIK----KVIONNESQSFVGTFTN 207
 Db 233 -TNQLTNTVTVGIDSGTTPHOAGYVKLVNFGFSVPNSAVKGDTFKITVTPKELNUNGVTST 291


```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-923-992A-2

Query Match 6.6%; Score 131.5; DB 4; Length 1164;
Best Local Similarity 22.8%; Pred. No. 0.15;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY 7 KKDANPNNGQTGLEA-ARMELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 59
Db 230 RKQQAQADKKEDAQVKREELGKLFSSTKAGLDQIEGHVKKETSSEENTQKVDEHYANS 289
QY 60 VNNNLNATLEQLMAKNLLESAINQANTDKTTTFONEHPNL-----VEAYKALKT 108
Db 290 LONLAQKSLLELDKATN-----EQATQVKNQFLENQAQKLEIQPLIKETNVKLYKAMSE 344
QY 109 TLEQATNLEGLSSTAYNGIRNNLVLDYLNKASSLITKLDPLNGGTTLLDNEIITANKNI 168
Db 345 SLEQVEKELKHNSA-----NLEDLVAKSKEIVREYEGKLNQSKNLP--ELKQLEEEA 395
QY 169 NNTLSTINEQ-----KTNADALSNSFFKKVI--QNNEQSFVGFTTANVQPSNYSFVAF 220
Db 396 HSKLKQVVEDFRKKFKTSEQVTPKKRVKRDLAANNENQOKI-----ELTVSPENI----- 445
QY 221 SADVTPVNYKARYARTVWNGDEPSSRILANTSITDVSWIYSLAGTNTKYOFSEF-----N 275
Db 446 -----TVTEGEDVKFTVTAKSQSKTTLDFSDLL-----TKYNPSVSDRISTN 487
QY 276 YGPSTG---XLYFPYKLVKAADANNVGLQYKLNNGNYQOVEFATSTSANNTTANP-TPAV 331
Db 488 YKTWNHKTAEIITIKNLNESQTVTLKAKDDSGNVVEKFTITVTQKKEQVPTPEQ 547
QY 332 DEIKVAKIV 340
Db 548 KDSKTEEKV 556

RESULT 13
US-08-923-992A-4
; Sequence 4, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707

```

```

; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-923-992A-4

Query Match 6.5% Score 131; DB 4; Length 1104;
Best Local Similarity 20.2%; Pred. No. 0.15; 154; Indels 132; Gaps 17;
Matches 85; Conservative 50; Mismatches 154; Indels 132; Gaps 17;

QY 7 KKDANPNNGQGLEA-ARMELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 59
DB 199 RKAQQAADKEDAQVYREELGKLFSSTKAGLQIQEHVKKTSSEBNTQKVDHEYANS 258
QY 60 VNNLNATLEOLKKAATNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 108
DB 259 LONLAQKSLEELDRAITN-----EQATQVRKQFLENQAQKLEIQPLIKETNVKLYKAMSE 313
QY 109 TLEQATNLEGLSTAVYQIRNNVLVDLYNKASSLITKTLDPINGTGLDLSNEITANKNI 168
DB 314 SLEQVEKELKHNSA-----NLQDLVAKSKEIVREYEGKLNQSKNLP---ELQKLEEEA 364
QY 169 NNTLSTINEQ-----KYNADALSNSFKKVI---ONNEQSVGFTFTNANVQPSYNSFVAF 220
DB 365 HSKLKQVVEYHPRKFKFTSEQVTPKKRVKRDIAAENNQOKI-----ELTVSPENI----- 414
QY 221 SADVTPVNYKARVTWNGDPPSSRIILANTNSITDVSNIYSLAGTNTKYQFS--FSNNGP 278
DB 415 -----TVYEGEDVKEFTVTAKSDS-----KTTLDFSDLLTKYNP 447
QY 279 STGYLYPPYKLVKAADANVGLQYKLLNGNVQVEFATSTSANNTANTPTPAVDEIKVAK 338
DB 448 SV-----SDRISTNYKTN-----TDNHKIAE 468
QY 339 IVLSGLREGON-TIELSPYPTGEGNM-----NKVAPMIGNIYLLSSNENNADKI 384
DB 469 ITINKLNQSQVTLLKAKDDSGNVVEKFTFTIVOKKEEQVQKPTPEQKHSKTEQNPQE 528
QY 385 P 385
DB 529 P 529

RESULT 14
US-09-268-347-48
; Sequence 48, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 2048
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-48

```

```

Query Match          6.4%; Score 129; DB 4; Length 2048;
Best Local Similarity 21.2%; Pred. No. 0.49;
Matches 83; Conservative 60; Mismatches 156; Indels 92; Gaps

QY  5  ITKKDAN-----PNNQGTQLEAARMELTDLINKAKMTLASLDQYAKIEASLSAYSEAE 58
      :|:|:|  ||  ||  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|
Db  288 VTQSDNNSRPAYTN---TOALDPKFQATNTRKAGPLSIGSNSIKRKI-INVGAGVNKTD 343
      :|:|:|  ||  ||  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|

QY  59 TVNNNLNATLEQKWKATNLESANQANTDKTTFDNEHPNLVYAYKALKTTLEQRATNLE 118
      ||  ||  ||  ||  :|:|:|:|:|:|  ||  ||  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|
Db  344 AVN-----VAQLEAVVYKWKARRITFGQD-DNSTQVKIGLDNLT-----TIK 384
      ||  ||  ||  ||  :|:|:|:|:|:|  ||  ||  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|

QY  119 GLSSTAYNOIRNNLVDLYNKA--SSLITKTLDPNGTGLDLSNEITTANKNNINTLSTN 176
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|
Db  385 GGAET--NALTDDNIGVYKADNSGLKVKLAKTLNNLT-----EVNTTLNATTTVKVGS 437
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|

QY  177 EQKTNADALSNSFIKKVIQNNQEOSFVGTFTNANVPQSNYSFVAFSADVTPVNYKYARRTV 236
      :|:|:|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db  438 SSSTTAELLSDSL-----TFQPN----- 456
      :|:|:|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY  237 WNGDEPSSRIILANTNSI--TDVSHIYSLAGTN--TKYQFSFNSYGPSTGYLYPPYKLKVA 292
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|
Db  457 -TGSQSTSKTVYGVYKFTNNAETTAAGTTTRITDKIGFARDG-DVDEKQAPYLDKKQ 514
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|

QY  293 ADANNVGLQKLNNGNVQVQEFATSTSANNTTANPTPAVDKIKVAKIVL-SGLRFGQNTI 351
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|
Db  515 LKVGSAI--TIDNG-IDAGNKKISLNKAGSSANDAVTIEQLKAARFTLNAGAGISVTTPT 571
      :|:|:|  :|:|:|:|:|:|  :|:|:|:|:|:|  ||  ||  :|:|:|:|:|:|  :|:|:|:|:|:|

QY  352 ELSVPTGEGNNKVPAMIGNIYSSNENNAD 382
      :|:|:|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db  572 EISVDAKSGNV--TAPTY-NIGVKTTELNSD 599
      :|:|:|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 15
US-08-923-992A-6
; Sequence 6, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1128 amino acids
; TYPE: amino acid

```


Search completed: June 12, 2002, 10:49:10
Job time: 203 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:48:38 ; Search time 55.68 Seconds
(without alignments)
785.975 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086
Perfect score: 2004
Sequence: 1 GCMSTFKDANPNNGQTGLE.....SSNENADKIPGRRPCTFL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2004	100.0	456	18 AAW36050	Hybrid Marek's dis
2	2000	99.8	1086	18 AAW36051	Hybrid Marek's dis
3	1910	95.3	615	15 AAR63230	Mycoplasma gallise
4	1856	92.6	610	15 AAR63229	Mycoplasma gallise
5	1612	80.4	368	14 AAR44493	Mycoplasma gallise
6	1612	80.4	368	15 AAR63227	Mycoplasma gallise
7	1604	80.0	368	16 AAR76955	Mycoplasma gallise
8	1142	57.0	235	10 AAP93646	Amino acid sequenc
9	1142	57.0	235	11 AAR05081	MG-1 antigen. AA
10	1142	57.0	261	11 AAR05082	TWG-1 antigen. A
11	1142	57.0	261	16 AAR79911	M.gallisepticum 26

12	1117	55.7	261	10 AAP93959	Amino acid (AA) se
13	806	40.2	661	15 AAR63226	Mycoplasma gallise
14	806	40.2	661	16 AAR79910	M.gallisepticum 66
15	713	35.6	647	16 AAW11978	Mycobacterium gall
16	700.5	35.0	648	15 AAR56973	PMGA 1.2 protein o
17	386.5	19.3	183	10 AAP93649	Amino acid sequenc
18	386.5	19.3	183	11 AAR06439	MG-4 antigen. AA
19	284.5	14.2	219	10 AAP93648	Amino acid sequenc
20	284.5	14.2	219	11 AAR06438	MG-3 antigen. AA
21	176.5	8.8	6281	22 AAU37403	Staphylococcus aur
22	176	8.8	1095	22 AAG83030	S. epidermidis ope
23	171	8.5	2086	22 AAU34143	Staphylococcus aur
24	171	8.5	5795	22 AAU37017	Staphylococcus aur
25	162	8.1	1029	22 AAU34389	Staphylococcus aur
26	162	8.1	1048	22 AAU37490	Staphylococcus aur
27	159	7.9	933	21 AAY58435	Staphylococcus aur
28	159	7.9	933	22 AAB69508	Staphylococcus aur
29	159	7.9	936	18 AAW89801	Staphylococcus aur
30	158	7.9	2434	22 AAU34339	Staphylococcus aur
31	156	7.8	807	21 AAB18311	Plasmodium falcipa
32	154.5	7.7	682	17 AAR95273	Nisin nisp gene pr
33	153.5	7.7	2478	22 AAU34320	Staphylococcus aur
34	153.5	7.6	2478	22 AAU37374	Staphylococcus aur
35	151.5	7.6	1237	18 AAW55640	H. pylori ORF 04ep
36	151.5	7.6	1237	20 AAY11787	H. pylori outer me
37	151	7.5	5024	22 AAG82935	S. epidermidis ope
38	145.5	7.3	2437	22 AAU34338	Staphylococcus aur
39	145	7.2	1215	22 AAU34412	Staphylococcus aur
40	145	7.2	1269	22 AAU37520	Staphylococcus aur
41	143.5	7.2	837	22 AAU34387	Staphylococcus aur
42	143.5	7.2	875	22 AAU37487	Staphylococcus aur
43	143	7.1	3158	22 AAU37018	Staphylococcus aur
44	142	7.1	135	15 AAR56974	Partial PMGA 1.3 p
45	142	7.1	135	16 AAW11979	Mycobacterium gall

ALIGNMENTS

RESULT	1
ID	AAW36050
AC	AAW36050 standard; Protein; 456 AA.
DT	15-JUL-1998 (first entry)
DE	Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.
KW	Chimeric; Marek's disease virus; outer membrane protein; fusion protein; antigen; vaccine; poultry.
OS	Chimeric - Marek's disease gammaherpesvirus.
FT	Key
FT	Region
FT	Location/Qualifiers
FT	1..64
FT	/note= "derived from Marek's disease virus gB protein"
FT	65..456
FT	/note= "derived from M. gallisepticum antigenic protein"
PN	WO9736924-A1.
PD	09-OCT-1997.
PF	28-MAR-1997; 97WO-JP01084.
PR	29-MAR-1996; 96JP-0103548.
PA	(JAPS) NIPPON ZEON KK.
PI	Saito S, Tsuzaki Y, Yanagida N;

DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 XX Disclosure; Page 16-19; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein GB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 XX Sequence 456 AA;
 SQ
 Query Match 100.0%; Score 2004; DB 18; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.3e-128;
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCMSTTKDANPNNGQQLAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATV 60
 Db 63 gcmsttkdanpnngqqlaarmeltdlinakamtiasldqdyakiesaisayseae 122
 Qy 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
 Db 123 nnnlnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnle 182
 Qy 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPNGGLTLDSEITANKNINNTLSTINEQKT 180
 Db 183 sstaynqirnnldvlynkasslittktdplnggtlldsneittankninnntlstineq 242
 Qy 181 NADALNSFIKKVIQNNQSFVGFFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 240
 Db 243 nadalsnfikkvqqneqsfvgtftnanvqpsnysfvaafadvtpvnykyartvwngd 302
 Qy 241 EPSSRILANTNSITDVSWIYSLAGTNTKYQFSNYSFGSTGYLYFPYKLVKAADANNVGL 300
 Db 303 epssrilantnsitdsvswiyslagntkyqfsnysfgstgylyfpykvlkaadannv 362
 Qy 301 QYKLNNGVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
 Db 363 qyklngnvqvvefatstsaanttanptpavdeikvakivlsgrlfrfgntiele 422
 Qy 361 NMNKVAPMIGNIYLSNENNADKIPYRRPCTFL 394
 Db 423 nmnkvapmigniylnsnennadkipgyrrpctfl 456
 RESULT 2
 ID AAW36051 standard; Protein; 1086 AA.
 XX
 AC AAW36051;
 XX
 XX 15-JUL-1998 (first entry)
 DT
 XX Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 DE
 XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 KW
 XX Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Region 1..672
 FT /note= "derived from Marek's disease virus GB protein"
 FT 693..1086
 FT /note= "derived from M. gallisepticum antigen"

XX WO9736924-A1.
 XX
 XX 09-OCT-1997.
 XX
 XX 28-MAR-1997; 97WO-JP01084.
 PF
 XX 29-MAR-1996; 96JP-0103548.
 PR
 XX (JAPG) NIPPON ZEON KK.
 XX
 XX Saito S, Tsuzaki Y, Yanagida N;
 PI
 XX WPI; 1997-503046/46.
 DR N-PSDB; AAT96596.
 DR
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 XX Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein GB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 XX Sequence 1086 AA;
 SQ
 Query Match 99.8%; Score 2000; DB 18; Length 1086;
 Best Local Similarity 99.7%; Pred. No. 7.5e-128;
 Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCMSTTKDANPNNGQQLAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATV 60
 Db 693 gcmsttkdanpnngqqlaarmeltdlinakamtiasldqdyakiesaisayseae 752
 Qy 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
 Db 753 nnnlnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnle 812
 Qy 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPNGGLTLDSEITANKNINNTLSTINEQKT 180
 Db 813 sstaynqirnnldvlynkasslittktdplnggtlldsneittankninnntlstineq 872
 Qy 181 NADALNSFIKKVIQNNQSFVGFFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 240
 Db 873 nadalsnfikkvqqneqsfvgtftnanvqpsnysfvaafadvtpvnykyartvwngd 932
 Qy 241 EPSSRILANTNSITDVSWIYSLAGTNTKYQFSNYSFGSTGYLYFPYKLVKAADANNVGL 300
 Db 933 epssrilantnsitdsvswiyslagntkyqfsnysfgstgylyfpykvlkaadannv 992
 Qy 301 QYKLNNGVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
 Db 993 qyklngnvqvvefatstsaanttanptpavdeikvakivlsgrlfrfgntiele 1052
 Qy 361 NMNKVAPMIGNIYLSNENNADKIPYRRPCTFL 394
 Db 1053 nmnkvapmigniylnsnennadkipgyrrpctfl 1086
 RESULT 3
 ID AAR63230 standard; Protein; 615 AA.
 XX
 AC AAR63230;
 XX
 XX 23-JUN-1995 (first entry)
 DT
 XX

DE Mycoplasma gallisepticum antigen (UM-67).
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
OS Mycoplasma gallisepticum.
FH Key Location/Qualifiers
FT Protein 1..615
FT /note= "Trp residues correspond to TGA codons"
XX
XX W09423019-A.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI: 1994-333181/41.
XX N-PSDB; AAQ77857.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 87-91; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX ORF encodes an antigenic polypeptide (AA63230). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 615 AA;

Query Match 95.3%; Score 1910; DB 15; Length 615;
Best Local Similarity 98.2%; Pred. No. 4.7e-122;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 CMSITKDKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 61
DB 27 cmsitkdkdanpnngqtlqearmeltdlinakartlasldyakiessaysaeatvn 86
QY 62 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 87 nninatleqlkmaiktlesainqantdkttfdnehpnlveaykalkttleqratinleqla 146
QY 122 STAYNQIRNNLVLYNKASSLIKTLDPLNGGTLIDSNEITTANKNNINNTLSINQKTN 181
DB 147 staynqirnnlvlynassliktldplnggmllidsneittvnrninntlstineqtn 206
QY 182 ADALSNFISFKKVTQNNBQSFVGTFTTANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
DB 207 adalsnfsfkvknqneqsfvgfttannvqpsnysfvafsadvtvpnykyartvwnge 266
QY 242 PSSRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 301
DB 267 psrillantnsitdvswiylagtntkyqfsnypstgylyfpyklykaadannvqlq 326
QY 302 YKLNGNVQVQFATSTSANNTANPTPAVDKIVKIVLSGLRFGQNTIELSVPTGEGN 361
DB 327 yklngnvqvqfatstsanntanptpavdkivkivlsgrfgqntielsvptgegn 386
QY 362 MNKVAPMIGNIYSSNENNADKI 384
|||||

Db 387 mnkvapmigniyssennennadki 409
RESULT 4
AAR63229
ID AAR63229 standard; Protein; 610 AA.
XX
XX AAR63229;
AC AAR63229;
XX
XX 23-JUN-1995 (first entry)
DT
XX
XX Mycoplasma gallisepticum antigen (UM-66).
DE
XX
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
KW
XX
XX Mycoplasma gallisepticum.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..610
FT /note= "Trp residues correspond to TGA codons"
FT
XX
XX W09423019-A.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI: 1994-333181/41.
XX N-PSDB; AAQ77856.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 610 AA;

Query Match 92.6%; Score 1856; DB 15; Length 610;
Best Local Similarity 95.5%; Pred. No. 2.2e-118;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
QY 2 CMSITKDKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 61
DB 27 cmsitkdkdanpnngqtlqearmeltdlinakamtiasldyakiessaysaeatvn 86
QY 62 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 87 nninatleqlkmaiktlesainqantdkttfdnehpnlveaykalkttleqratinleqls 146
QY 122 STAYNQIRNNLVLYNKASSLIKTLDPLNGGTLIDSNEITTANKNNINNTLSINQKTN 181
DB 147 staynqirnnlvlynkassliktldplnggtlidsneittanknnintlstineqtn 206
QY 182 ADALSNFISFKKVTQNNBQSFVGTFTTANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
DB 207 adalsnfsfkvknqneqsfvgfttannvqpsnysfvafsadvtvpnykyartvwnge 266

QY 242 PSSRILANTNSITDVSIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLVKAAADANNVGLQ 301
 |||||
 Db 267 pssrilantnsitdsvsiysslgtntkyqfsfsnygstgylyfpyklvkaadasnvgllq 326
 |||||
 QY 302 YKLNNGNVQVEFATSNANTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 361
 |||||
 Db 327 yklnngnvqvefatssannttanptpavdeikvakivlsgrlfrgntielsvptgern 386
 |||||
 QY 362 MNKVAPMIGNIYSSNENNADK 383
 |||||
 Db 387 mnkvapmignmyitssnaeank 408
 |||||
 RESULT 5
 AAR44493
 ID AAR44493 standard; Protein; 368 AA.
 AC AAR44493;
 XX
 DT 16-JUN-1994 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW Vaccine; mycoplasma infection; poultry; fowl.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262
 FT /note= "corresponds to NNN codon in AAQ53419"
 FT Misc-difference 283
 FT /note= "corresponds to NNN codon in AAQ53419"
 FT
 XX
 XX WO9324646-A.
 XX
 XX 09-DEC-1993.
 XX
 XX 28-MAY-1993; 93WO-JP00715.
 XX
 XX 29-MAY-1992; 92JP-0138819.
 XX
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
 XX
 XX WPI; 1993-405837/50.
 DR N-PSDB; AAQ53419.
 XX
 XX Mycoplasma gallisepticum antigen and DNA coding for it - useful
 PT for vaccination of fowl against mycoplasma infections
 XX
 PS Claim 2; Page 23-26; 37pp; Japanese.
 XX
 CC The sequence coding for the 40kDa antigen was obtained by PCR
 CC amplification of M.gallisepticum genomic DNA. The antigen reacts
 CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used
 CC as a vaccine to protect fowl from M.gallisepticum infection.
 XX
 XX Sequence 368 AA;
 SQ

Query Match 80.4%; Score 1612; DB 14; Length 368;
 Best Local Similarity 95.0%; Pred. No. 4.8e-102;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 QY 2 CMSITKDKANPNNGTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61
 |||||
 Db 27 cmsitkdkanpnngtqlearmeltdlinakartilasldyakielaslsayseaeetvn 86
 |||||
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 |||||

Db 87 nnlnatleqlkmaaktulesainganctdktfdnehpnlveaykalkttleqratnlegla 146
 QY 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTANKNINNTLSTINEQKTN 181
 |||||
 Db 147 staynqirnnldvlynasslittktdplngmldnsneittvnrnintlstinegktn 206
 |||||
 QY 182 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWNNGDE 241
 |||||
 Db 207 adalsnsfikkvionneqsfvgtftnanvqpsnysfafsadvtpvnykyarrtvxngde 266
 |||||
 QY 242 PSSRILANTNSITDVSIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLVKAAADANNVGLQ 301
 |||||
 Db 267 pssrilantnsitdsvsiysslgtntkyqfsfsnygstgylyfpyklvkaadannvgllq 326
 |||||
 QY 302 YKLNNGNVQVEFATSNANTTANPTPAVDEIKVAK 338
 |||||
 Db 327 yklnngnvqvefatssannttanptqgimrklk 363
 |||||
 RESULT 6
 AAR63227
 ID AAR63227 standard; Protein; 368 AA.
 XX
 AC AAR63227;
 XX
 DT 23-JUN-1995 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW recombinant avipox virus; live vaccine; mycoplasma 40kd antigen;
 KW TTM-1.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262
 FT /note= "corresponds to a NNN codon"
 FT Misc-difference 283
 FT /note= "corresponds to a NNN codon"
 FT
 XX
 XX WO9423019-A.
 XX
 XX 13-OCT-1994.
 XX
 XX 31-MAR-1994; 94WO-JP00541.
 XX
 XX 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeki S, Saitos, Takahashi K;
 XX
 XX WPI; 1994-333181/41.
 DR N-PSDB; AAQ77854.
 XX
 XX Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 PT of a live vaccine
 XX
 PS Claim 4; Page 71-74; 123pp; Japanese.
 XX
 CC The plasmid pUTTM-1P contains a sequence (the TTM-1 gene) coding
 CC for the 40kD antigen of Mycoplasma gallisepticum under the control
 CC of a synthetic promoter. A 1300 bp restriction fragment containing
 CC the promoter-ORF sequence was excised and was used in the
 CC construction of plasmid pNZ7929-R2. This in turn was involved in the
 CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FpV
 CC sequences. The recombinant avipox virus is useful as a live vaccine

XX This amino acid sequence of MG1 is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
 XX
 SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 10; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKKDANPNNGQTOLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61
 DB 1 cmsitkkdanpnngqtlqarmeltdlinakartlasldyakiesslsayseaeatvn 60
 QY 62 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 DB 61 nninatleqlkmaaktlnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 120
 QY 122 STAYNQIRNNLVLYNKASSLIKTLDPLNGGLDLSNEITTANKNNINNTLSTINEQKTN 181
 DB 121 staynqirnnlvlynassliktldplnggmlldsneittvnrnnntlstineqktn 180
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
 DB 181 adalsnfsikkvionneqsfvgtftnanvqpsnysfvafsadvtpvnykyarrtv 235

RESULT 9

AA05081
 ID AAR05081 standard; protein; 235 AA.
 AC AAR05081;
 XX
 DT 08-OCT-1990 (first entry)
 XX
 DE MG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX

JP02111795-A.

24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX

{JAPG } NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX

WPI; 1990-169109/22.
 DR N-PSDB; AAQ04686.
 XX

XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX

PS Claim 2; Fig 1a; 20pp; Japanese.
 XX

CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081-2 and AAR06437-41.
 XX

SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 11; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKKDANPNNGQTOLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61
 DB 1 cmsitkkdanpnngqtlqarmeltdlinakartlasldyakiesslsayseaeatvn 60
 QY 62 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 DB 61 nninatleqlkmaaktlnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 120
 QY 122 STAYNQIRNNLVLYNKASSLIKTLDPLNGGLDLSNEITTANKNNINNTLSTINEQKTN 181
 DB 121 staynqirnnlvlynassliktldplnggmlldsneittvnrnnntlstineqktn 180
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
 DB 181 adalsnfsikkvionneqsfvgtftnanvqpsnysfvafsadvtpvnykyarrtv 235

RESULT 10

AA05082
 ID AAR05082 standard; protein; 261 AA.
 XX

AC AAR05082;
 XX

DT 08-OCT-1990 (first entry)
 XX

DE TMG-1 antigen.
 XX

XX Mycoplasma gallisepticum; poultry; vaccine.
 XX

PN JP02111795-A.
 XX

PD 24-APR-1990.
 XX

PF 02-JUN-1989; 89JP-0136343.
 XX

PR 02-JUN-1989; 89JP-0136343.
 XX

PA {JAPG } NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX

WPI; 1990-169109/22.
 DR N-PSDB; AAQ04687.
 XX

XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX

PS Claim 2; Fig 2; 20pp; Japanese.
 XX

CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.
 XX

SQ Sequence 261 AA;

Query Match 57.0%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 3.2e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKKDANPNNGQTOLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 61
 DB 27 cmsitkkdanpnngqtlqarmeltdlinakartlasldyakiesslsayseaeatvn 86
 QY 62 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 DB 87 nninatleqlkmaaktlnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 146
 QY 122 STAYNQIRNNLVLYNKASSLIKTLDPLNGGLDLSNEITTANKNNINNTLSTINEQKTN 181


```
|||||
Db 147 staynqirnnldvlynnasslltktdlplngmllsneittvnrnnntlstineqkcn 206
|||||
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
|||||
Db 207 adalnsfikkvigneqsfvgtftnnavqpsnysfvafsadvtpvnykyarrtv 261
|||||

RESULT 11
AAR79911
ID AAR79911 standard; Protein; 261 AA.
XX
AC AAR79911;
XX
DT 19-JUL-1996 (first entry)
XX
DE M.gallisepticum 261 amino acid protein.
XX
KW Detection; probe; primer; PCR; amplification; secretion; lung;
KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX
OS Mycoplasma gallisepticum.
XX
PN JP07236498-A.
XX
PD 12-SEP-1995.
XX
PF 25-FEB-1994; 94JP-0052764.
XX
PR 25-FEB-1994; 94JP-0052764.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
WPI; 1995-347462/45.
DR N-PSDB; AAT04076.
XX
XX
XX Detection of Mycoplasma gallisepticum - for the quick detection,
PT i.e. within one day, of avian chronic respiratory diseases
XX
PS Claim 3; Page 10-11; 11pp; Japanese.
XX
XX This is the amino acid sequence of a 261 amino acid protein encoded
CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
CC sequence and the sequence of AAT04075 (encoding a 661 amino acid
CC protein) can be used to detect M.gallisepticum using probes based on
CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
CC complement of bases 893-919, 1908-1934 and the complement of bases
CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
CC encoding this protein. The method using these sequences is faster i.e. is
CC able to detect M.gallisepticum, which causes avian chronic respiratory
CC diseases, within one day, from avian secretions, washings from the lung,
CC respiratory tract, nasal cavity, etc.
XX
SQ Sequence 261 AA;

Query Match 57.0%; Score 1142; DB 16; Length 261;
Best Local Similarity 97.0%; Pred. No. 3.2e-70;
Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
|||||
Db 27 cmsitkddanpnngqtqlqaarmeltdlinakartiasldyakielasaysaeatvn 86
|||||
Qy 62 NNUNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVLYEAYKALKTTLEQATNLEGLS 121
|||||
Db 87 nnlnatleqlkmaktulesainqandctktdfndehnpvlyeaykalkttleqatnlegla 146
|||||
Qy 122 STAYNQIRNNLDVLYNNKASSLTKTLDPLNGGTLDSNEITTANKNNNTLSTINEQKTN 181
|||||
Db 147 staynqirnnldvlynnasslltktdlplngmllsneittvnrnnntlstineqkcn 206
|||||
```

```
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
|||||
Db 207 adalnsfikkvigneqsfvgtftnnavqpsnysfvafsadvtpvnykyarrtv 261
|||||

RESULT 12
AAP93959
ID AAP93959 standard; protein; 261 AA.
XX
AC AAP93959;
XX
DT 11-MAY-1990 (first entry)
XX
DE Amino acid (AA) sequence of TMG-1 polypeptide.
XX
KW Mycoplasma gallisepticum; Poultry vaccine; ss;
XX
OS Mycoplasma gallisepticum.
XX
PN EP345021-A.
XX
PD 06-DEC-1989.
XX
PF 31-MAY-1989; 89EP-0005441.
XX
PR 02-JUN-1988; 88JP-0136343.
XX
PA (JAPG ) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.
XX
PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
XX WPI; 1989-358393/49.
DR N-PSDB; AAN92574.
XX
XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT vaccines.
XX
PS Disclosure; Fig.2; 31pp; English.
XX
XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
CC same sequence as that of a polypeptide expressed in Mycoplasma
CC gallisepticum in nature. When the corresponding DNA sequence is inserted
CC into a recombinant vector used to transform a host the antigen protein
CC produced can be used as a vaccine to prevent and diagnose MG infection.
XX
SQ Sequence 261 AA;

Query Match 55.7%; Score 1117; DB 10; Length 261;
Best Local Similarity 96.2%; Pred. No. 1.6e-68;
Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;

Qy 2 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
|||||
Db 27 cmsitkddanpnngqtqlqaarmeltdlinakartiasldyakielasaysaeatvn 86
|||||
Qy 62 NNUNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVLYEAYKALKTTLEQATNLEGLS 121
|||||
Db 87 nnlnatleqlkmaktulesainqandctktdfndehnpvlyeaykalkttleqatnlegla 146
|||||
Qy 122 STAYNQIRNNLDVLYNNKASSLTKTLDPLNGGTLDSNEITTANKNNNTLSTINEQKTN 180
|||||
Db 147 staynqirnnldvlynnasslltktdlplngmllsneittvnrnnntlstineqkkt 205
|||||
Qy 181 NADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
|||||
Db 206 nadalnsfikkvigneqsfvgtftnnavqpsnysfvafsadvtpvnykyarrtv 261
|||||

RESULT 13
AAR63226
ID AAR63226 standard; Protein; 661 AA.
XX
```

```

AC AAR63226;
XX
DT 23-JUN-1995 (first entry)
XX
DE Mycoplasma gallisepticum antigen (UM-81).
XX
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Protein 1..661
FT /note= "Trp residues correspond to TGA codons"
XX
PN W09423019-A.
XX
PD 13-OCT-1994.
XX
PF 31-MAR-1994; 94WO-JP00541.
XX
PR 31-MAR-1993; 93JP-0074139.
PR 30-SEP-1993; 93JP-0245625.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
PI Saeki S, Saitos, Takahashi K;
XX
DR WPI; 1994-333181/41.
DR N-PSDB; AAQ77853.
XX
PT Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PT of a live vaccine
XX
PS Claim 4; Page 61-65; 123pp; Japanese.
XX
CC A restriction fragment of the insert of M.gallisepticum genomic clone
CC pUM-81 containing an open reading frame was sequenced (AAQ77853). The
CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox
CC virus comprising the coding sequence can be used as a live vaccine to
CC protect against infection by Mycoplasma gallisepticum.
XX
SQ Sequence 661 AA;

Query Match 40.2%; Score 806; DB 15; Length 661;
Best Local Similarity 42.8%; Pred. No. 8.8e-47;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

QY 9 DANPNNGQTQLEAARMELTD-----LINAKAMTLASLDQYAKIEASLSAYSEATVN 61
DB :||:||||: || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45 ntspdsdgmmnaaakeladakaalttlingetanlasedyakikseltsayetakavs 104

QY 62 NNLNATLEOLKMATNLSEALNQANTDKTTFDNEHPNLVEAYKALKTLEQRATNLEGLS 121
DB ||||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 aktgatneveakttdaalkaasakndfdaghslveaynnlketlkeektnldsla 164

QY 122 STAYNQIRNNLVLDYLNKASSLITTTDPLNGGTTLLDSNEITANKNNNTLSTINEQKTN 181
DB : || || || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 nenyaairthlnslyekantivatltpat-gnipewmsvtcgaqgditnatsrliawkqn 223

QY 182 ADALNSNFIKKVIQNQEQSGYGTFTNANVQPSNYSFVAFSADVTFPVNYKYARRTVWNGDE 241
DB ||||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 adnlansfikqlvknlnitrvdvanqqe-qpanysfvgfsvndtpwnfnqkrkwasen 282

QY 242 -----PSSRLNANVNSTIDVSWIYSIAGTNTKYQFSPSNYG-PSTGYLYFPYKLVRAAD 294
DB : ||||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 tplattpaedatqqaaltdvswiysingaekytlslfryfgaektylyfpyklvktstd 342

QY 295 ANNYGLOYKLNNGNVQOQVEF-----ATSTPSAN-NTTANP-----TPAVDEIKVAKIV 340
DB ||||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 343 --nvglqykingdtkqinfvqtpasgssdvaaneetmaspaemqsaaptvddikiakva 400
QY 341 LSGLRFGONTIELSVPTGEGNMKNKVPIMIGNIYLSS-----NENN 380
DB ||||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 lsnlkfsntiefsvptg-----kaapmignmyltsnsevnknklyddlfgnsfnenn 455
QY 381 -----ADKIPGY 387
DB 456 ptavtvdllkg 467

RESULT 14
AAR79910
ID AAR79910 standard; Protein; 661 AA.
XX
AC AAR79910;
XX
DT 19-JUL-1996 (first entry)
XX
DE M.gallisepticum 661 amino acid protein.
XX
KW Detection; probe; primer; PCR; amplification; secretion; lung;
KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Misc-difference 220
FT /note= "encoded by TGA"
FT Misc-difference 270
FT /note= "encoded by TGA"
FT Misc-difference 305
FT /note= "encoded by TGA"
FT Misc-difference 581
FT /note= "encoded by TGA"
XX
PN JP07236498-A.
XX
PD 12-SEP-1995.
XX
PF 25-FEB-1994; 94JP-0052764.
XX
PR 25-FEB-1994; 94JP-0052764.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
WPI; 1995-347462/45.
N-PSDB; AAT04075.
XX
Detection of Mycoplasma gallisepticum - for the quick detection,
i.e. within one day, of avian chronic respiratory diseases
Claim 2; Page 7-10; 11pp; Japanese.
XX
This is the amino acid sequence of a 661 amino acid protein encoded
by a fragment of the Mycoplasma gallisepticum genome. The encoding
sequence and the sequence of AAT04076 (encoding a 261 amino acid
protein) can be used to detect M.gallisepticum using probes based on
nucleotides 1125-1648 and primers based on nucleotides 449-466, the
complement of bases 893-919, 1908-1934 and the complement of bases
2184-2210 of the sequence, and a probe based on nucleotides 718-41 of
AAT04076. The method using these sequences is faster i.e. is able to
detect M.gallisepticum, which causes avian chronic respiratory diseases,
within one day, from avian secretions, washings from the lung,
respiratory tract, nasal cavity, etc.
XX
SQ Sequence 661 AA;

Query Match 40.2%; Score 806; DB 16; Length 661;
Best Local Similarity 42.8%; Pred. No. 8.8e-47;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

```

QY	9	DANPNNGOTOLEAARMELTD-----LINAKAMTILASLODYAKIEASLSSAYSAPETVN	51
		: : : : : : : : : : : : : :	
Db	45	ntupsdgqgmnaaaakeladakaaltllingetancilasiedyakikseltsayetakavs	104
		: : : : : : : : : : : : : :	
QY	62	NNUNATLEQLKMAKTLESAINOANTDKTTFDNEHPNLVAYKALKTTTLEQRATNLEGIS	121
		: : : : : : : : : : : :	
Db	105	aktgatinevneaktldaaikkaasakndfaqhslyeayonlketlikeektnldsia	164
		: : : : : : : : : : : : : :	
QY	122	STAYNOIRNNLDVLYKASSLIPTKLPLUNGTLTLLDSNETITANKNNINTLTINEOKTN	181
		: : : : : : : : : : : : : :	
Db	165	nennyairtinsilyekanvicatldpat-gnipemsvtgangdoitnatrlawkqn	223
		: : : : : : : : : : : : : :	
QY	182	ADALSNSFIKKVIIONNEQSFVGFTFNANYQPNSIFYAFSADVTPVNYKYARTRVWNGDE	241
		: : : : : : : : : : : :	
Db	224	adnlansfikqsivknltrvdvanqe-qpanysfvgfsvnvdtpnwdaqrkvwaseen	282
		: : : : : : : : : : : : : :	
QY	242	-----PSSRLIANTNSTDVSIYSLAGTNKYQFSPSNYG-PSGUYLYFPYKLVKAAD	294
		: : : : : : : : : :	
Db	283	tplattpaedatqaasiltdvswyisingaeakyltsrfygaehtaylyfpfkvlktkd	342
		: : : : : : : : : : : : : :	
QY	295	ANNVGLQYLKNNNVQOVEF-----ATSTSAN--NTTANP-----TPAVDEIKVAKIV	340
		: : : : : : : : : : : : : : : : : :	
Db	343	--nvglqyklinggtkginfvtqpasgssdvaaneesetmaspaemqsaaptvddiktakwa	400
		: : : : : : : : : : : : : :	
QY	341	LSGLRSCONTIELSVPTGECNNKNVAPMIGNTYLSS-----NENN	380
		: : : : : : : : : : : : :	
Db	401	lsnuknsnctiefsvptg-----kaapmigmnyiltssnevknklyddlfngsfnnenn	455
		: : : : : : : : : : : : : :	
QY	381	-----ADKIPGY	387
		: : : : : : : : : : : : : :	
Db	456	ptavtvdlilkgy	467

RESULT 15
AAW11978
ID AAW11978 standard; Protein; 647 AA.

AC AAW11978;

XX 5

DT 23-APR-1955 XX

DE Mycobacter:

XX

KW Adhesin; P
KW Adhesin; PXX
KW

OS Mycoplasma

XXII

FH	Key
FT	pentide

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
8

LE

FT YY

AA
PN CA2135330-

XX

PD 11-MAY-1961

XX
PF
08-NOV-196

XX

PR 20-APR-1991

XX
PR 10-NOV-01 16T-AON-01

PA (BROW) BR

XX

XX
PI Browning &

DR WPI; 1995

DR N-PSDB; A

XX
DT
New promo

PT New promo
PT useful wh

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
8

PT	
multivalent live vaccines, also new probes for detecting Mycoplasma	
PT	
and manipulating its genome	
XX	
Disclosure; Fig 3; 8ipp: English.	
XX	
Adhesin pmGAL.2 (AAW11978) and adhesin pmGAL.3 fragment (AAW11979)	
CC	
are products of gene sequences (see also AAT51531) isolated from	
CC	
Mycoplasma gallisepticum. DNA constructs incorporating the	
CC	
promoter and/or signal sequences of the pmGA genes can be used	
CC	
in the prodn. of multivalent live vaccines. The signal peptide	
CC	
sequence is utilised where attachment of an exogenous antigen	
CC	
gene to the mycoplasma cell membrane is required.	
XX	
Sequence 647 AA:	
SQ	

Query Match 35.6%; Score 713; DB 16; Length 647;
Best Local Similarity 41.1%; Pred. NO. 1.9e-40;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps

QY	9	DANPNQC-----TQLEAARMELTDLINAKAMTILASLODYAKIEBASUSSAYSEATVN	61
Db	48	dt:pnpgdgmmnaasgelaarmglttifgskaknlglyvdykktqntlckaydaaktvl	107
QY	62	NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTILEQRAFNLGLS	121
Db	108	dnssttqklnearktrletairtaatskqtfdeghaelvkvkykelktltsnetatlapy	167
QY	122	STAYNQIRNNLDVLINAKASSILKTLIDPLNGGTLILDSNETTTANKNNINTL--STINEOK	179
Db	168	dagygagikmhlglydagkaitktiepvsgdp-ltasavmmanatkiveaidevinpqk	226
QY	180	TN-----ADALSNSEFIKKVIONNQSFVGTGTNNANVOPSNYSFVAFSADVTPV-----	227
Db	227	enatk1adslsllsvkkitvee-----ahnkadpanysfvgykrywtellldkqvfp	279
QY	228	NYKARTVW-NGDEPSSRLANT-----NSITDVSWISYSLAGTNKYQSFSGNYGP	278
Db	280	nwdyaertiftnsdep--rsisntpadgqtnaaplnsvswisylagtakytlefkyfyp	337
QY	279	STGYLXFPYKLVKAADANNYGLAYKLNMGNVQOFEFATSTSA-----NNTNANTPTPAVDE	333
Db	338	stgylfpyk1vntsdqvk1gleyk1nd-----atkpasaitfgsdqtmngktpvtnd	399
QY	334	IKVAKIVLSGERFGQNTIELSVPTGEGNNKVPMPMIGNIYLSNENNAADKI	384
Db	390	invakvlelanlfgnsk1iefsva-----ekvsgpm1gmymylsspspnwnk1	435

Search completed: June 12, 2002, 10:48:39
Job time: 387 sec

FT	Peptide	1..25	location/coordinates
FT		/label= Sig.peptide	
FT		/note= "the signal peptide shows homology to	
FT		the pmGal.3 signal peptide"	
XX			
XX			
PN	CA2135330-A.		
XX			
PD	11-MAY-1995.		
XX			
XX	08-NOV-1994;	94CA-2135330.	
XX			
PR	20-APR-1994;	94US-0230312.	
PR	10-NOV-1993;	93AU-0050593.	
XX			
XX	(BROW/) BROWING G F.		
XX			
PI	Browning GF, Glew MD, Markham PF, Walker ID, Whithear KG;		
XX			
DR	WPI; 1995-241027/32.		
DR	N-PSDB; AAT51531.		
XX			
XX			
PT	New promoter region from a Mycoplasma gallisepticum adhesin gene -		
PT	useful when coupled to foreign antigen gene, for prodn. of		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:25:13 ; Search time 98 Seconds
(without alignments)
1037.477 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086
Perfect score: 2004
Sequence: 1 GCMSTTKDANPNNGQTGLE.....SSNENNADKIPYRRPGTFL 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rident.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	59.9	271	2 Q49464	Q49464 mycoplasma
2	837	41.8	671	2 Q9L8D6	Q9L8D6 mycoplasma
3	831.5	41.5	702	2 Q49499	Q49499 mycoplasma
4	826	41.2	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	40.2	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	38.0	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	37.6	702	2 Q49498	Q49498 mycoplasma
8	746	37.2	645	2 Q9Z1D1	Q9Z1D1 mycoplasma
9	745	37.2	644	2 Q9ZHR9	Q9ZHR9 mycoplasma
10	744	37.1	650	2 Q49495	Q49495 mycoplasma
11	743	37.1	649	2 Q49497	Q49497 mycoplasma
12	739.5	36.9	486	2 Q9RLX9	Q9RLX9 mycoplasma
13	713	35.6	647	2 Q49468	Q49468 mycoplasma
14	704	35.1	656	2 Q9KH15	Q9KH15 mycoplasma
15	690.5	34.5	703	2 Q05122	Q05122 mycoplasma
16	619.5	30.9	419	2 Q9L8D5	Q9L8D5 mycoplasma

17	575.5	28.7	386	2 Q49500	Q49500 mycoplasma
18	557	27.8	367	2 Q9XCG7	Q9XCG7 mycoplasma
19	189.5	9.5	320	2 Q49496	Q49496 mycoplasma
20	175	8.7	6713	16 Q99U54	Q99U54 staphylococ
21	175	8.7	6713	16 Q931R6	Q931R6 staphylococ
22	175	8.7	9439	16 Q8CP76	Q8CP76 staphylococ
23	173	8.6	1302	2 Q49547	Q49547 mycoplasma
24	170	8.5	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
25	168.5	8.4	810	5 Q8I4Y6	Q8I4Y6 plasmodium
26	165.5	8.3	661	16 Q8YUJ8	Q8YUJ8 anabaena sp
27	164.5	8.2	2481	16 Q99Q86	Q99Q86 staphylococ
28	163.5	8.2	3890	16 Q99U53	Q99U53 staphylococ
29	160.5	8.0	433	16 Q97TE6	Q97TE6 clostridium
30	160.5	8.0	682	2 Q48674	Q48674 lactococcus
31	159.5	8.0	1631	5 Q8I3H3	Q8I3H3 plasmodium
32	159	7.9	933	2 Q53653	Q53653 staphylococ
33	159	7.9	1774	5 Q8IAU8	Q8IAU8 plasmodium
34	158.5	7.9	946	16 Q8NXJ1	Q8NXJ1 staphylococ
35	158.5	7.9	1344	2 Q49545	Q49545 mycoplasma
36	157.5	7.9	3001	5 Q8IL08	Q8IL08 plasmodium
37	156	7.8	1149	5 Q9G263	Q9G263 plasmodium
38	155.5	7.8	3178	5 Q8I2P8	Q8I2P8 plasmodium
39	155	7.7	1072	16 Q9CF64	Q9CF64 lactococcus
40	155	7.7	1461	16 Q8ZL64	Q8ZL64 salmonella
41	153	7.7	4688	16 Q9PQ08	Q9PQ08 ureaplasma
42	154.5	7.7	1365	2 Q49525	Q49525 mycoplasma
43	154	7.7	1979	5 Q8ICX4	Q8ICX4 plasmodium
44	153.5	7.7	1795	16 Q9LCJ9	Q9LCJ9 staphylococ
45	153.5	7.7	2478	2 Q9RL69	Q9RL69 staphylococ

ALIGNMENTS

RESULT 1

Q49464	Q49464	PRELIMINARY;	PRT;	271 AA.
ID	Q49464			
AC	Q49464;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	TM-1 (Fragment).			
GN	TM-1.			
OS	Mycoplasma gallisepticum.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_TaxID=2096;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94025893; PubMed=8212828;			
RA	Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,			
RA	Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.;			
RT	"Cloning and DNA sequence of a 29 kilodalton polypeptide gene of			
RT	Mycoplasma gallisepticum as a possible protective antigen.";			
RL	Vaccine 11:1061-1066(1993).			
DR	EMBL; S65869; AAB28343.2; -.			
FT	NON_TER 271			
SQ	SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;			

Query Match 59.9%; Score 1200; DB 2; Length 271;
Best Local Similarity 97.1%; Pred. No. 1.5e-49;
Matches 238; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY	2	CMSTTKDANPNNGQTGLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSAETVN	61
DB	27	CMSTTKDANPNNGQTGLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSAETVN	86
QY	62	NNLNATLEQLKMAKTNLESAINQANTDKTTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS	121
DB	87	NNLNATLEQLKMAKTNLESAINQANTDKTTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA	146
QY	122	STAYNQIRNRLVDLTKKASSLITKTLDPNGTLTDSNEITTANKNINNTLSTINEQKTN	181
DB	147	STAYNQIRNRLVDLTKKASSLITKTLDPNGTLTDSNEITTANKNINNTLSTINEQKTN	206

```

QY 182 ADALSNFISFKKVIQNNQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTYWNGDE 241
DB 207 ADALSNFISFKKVIQNNQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTYWNGDE 266
QY 242 PSSRI 246
DB 267 PSSRI 271

RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
AC Q9L8D6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PMGA-like protein 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN SEQUENCE FROM N.A.
RP STRAIN=F;
RC PHARR G.T., Branton S.L., Hanson L.A., Minlon F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum."
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAD29524.1; -.
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;

Query Match 41.8%; Score 837; DB 2; Length 671;
Best Local Similarity 43.4%; Pred. No. 5e-32;
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;

QY 9 DANP-----NGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVNN 62
DB 55 DTNFGNDGGMNSAQQAQAAKELSDLLATQSSNLAKYADYNTIONTLTAAYTTAKSTD 114
QY 63 NLNATLEQKMAKTNLESAINQANTDKTTFDNEHNLVEAYKALKTTLEQATNLEGLSS 122
DB 115 NTSVLEQVKSATSTLQAADTAASSKTSFDEKPNELIKAYALKTKNEETVLSGLTD 174
QY 123 TAYNOIRNLDVLYNKASSLTITKLDPLNGGTLDSNEITANKNINNTLSTINEQNTNA 182
DB 175 SNFATIKTNLTALYQSGKDFVKATLDLPVSGNA-PQIADITKADKDIADAVSKLETWKTNA 233
QY 183 DALSNFTKVIQNNQSFVGTFTNANQPSNYSFVAFSADVTP-----VNYKYARRTYWN 238
DB 234 NTLATSFVKELVKNKLTGIDT-TNNRQPGNYSEVGYSVNATNNNEIPNNFAQRKVV 292
QY 239 GDEPSSRILA-----NTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 294
DB 293 SDNGRTSLISDTSNDSSTLTVEVSWIYSLGAGTKYSLTFNYGPGSTGYLYFPYKLVK 352
QY 295 ANNVGLQYKLVNNGVQVEFA-----TSTSA-----NNTT-----ANPTPAVD 332
DB 353 ENNVALQYTLNLSGSAQVNFAPTKVTSVSDSSGDSNNQTESAAETMPVTDLPAPTVS 412
QY 333 EIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADKI 384
DB 413 DINIAKTLNLSLKFSGNTIEFSVPTPEPS--NKVAPMIGNMYLTSNIANEAKV 462

RESULT 3
Q49499 PRELIMINARY; PRT; 702 AA.
AC Q49499;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGA1.4 protein precursor.

GN PMGA1.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN SEQUENCE FROM N.A.
RP STRAIN=S6;
RC MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62418.1; -.
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;

Query Match 41.5%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 9.6e-32;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 9 DANPNG-----QTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVNN 62
DB 68 NTPNGNGGTDNAQAQAAKELSDLLATQSSNLSTVADYANIQTTLTAAYTTAKSTD 127
QY 63 NLNATLEQKMAKTNLESAINQANTDKTTFDNEHNLVEAYKALKTTLEQATNLEGLSS 122
DB 128 NTSATLEQVKSATSTLQTAIDTAASSKTSFDEKPNELIKAYALKTKWNRNSUSGLTD 187
QY 123 TAYNOIRNLDVLYNKASSLTITKLDPLNGGTLDSNEITANKNINNTLSTINEQNTNA 182
DB 188 SNFATIKTNLTALYQSGKDIVTKLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTNA 246
QY 183 DALSNFTKVIQNNQSFVGTFTNANQPSNYSFVAFSADVTP-----VNYKYARRTY 236
DB 247 TVLATSFVKELVKNKLTGIDT-TNNRQPGNYSEVGYSDVDTGSDNARNPNWSEAFQKV 305
QY 237 WNGD-----EPSSRIILANTNITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLV 291
DB 306 WTSNDILSQPAEENQSQSAPDVSWIYNLGNGAKYSLTFNYGPGSTGYLYFPYKLV 365
QY 292 AADANNVGLQYKLVNNGVQVEFATS-----TSANN-----TTANPTPAVD 332
DB 366 SSDSDKVALEYKLNEASVKTIDFSPQTSVPASDATRENNRSTAAPAOQSGSTEINPAPTLD 425
QY 333 EIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADK 383
DB 426 DIKIAKTVLSLKFSGNTIEFSVPTTAKEGTSKVAPMIGNMYLTSNDRDVK 477

RESULT 4
Q9XCG8 PRELIMINARY; PRT; 632 AA.
AC Q9XCG8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VLH1 precursor (Fragment).
DE VLH1.
GN Mycoplasma imitans.
OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=29560;
RN SEQUENCE FROM N.A.
RP STRAIN=4229;
RC MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Grew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans
RT family of Mycoplasma gallisepticum."
RL Microbiology 145:2095-2103(1999).
DR EMBL; AF141940; AAD39483.1; -.
KW Signal.
```

```

FT NON_TER 1 1
FT SIGNAL <1 17 POTENTIAL.
FT CHAIN 18 632 VLHAL.
SQ SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;

Query Match
Best Local Similarity 41.2%; Score 826; DB 2; Length 632;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

QY 8 KDANPNNGOT-----QLEAARMELTDLINAKAMTSLASQDYAKTEASLSAYSAE 58
DQ 34 KPNPDTGNTNPGDGTDAKQADAKASLNTLLGSOQTNVYALYAKIKDILSSAYSAQ 93
QY 59 TVNNNLATLEOLKMAKNLESAINQANTDKTTTDENEHPNLVEAYKALKTTLEQATNLE 118
DQ 94 TTANNANATLVNDAAKALQATASDAANAKTEFDKANGLVSAAYAKIKETLKSETTND 153
QY 119 GLSTAYNOIRNNVLNLYNKASSLITKTLDPLNGTLLDSNEITTANKNNINLTSTINEQ 178
DQ 154 GLSQDNYSAIKANLVSLYNKAKDFTTVDLDP-TSGMIPKVDIEITSANTAITQAVSAIDQ 212
QY 179 KTNADALNSFIKKVIONNEQSFVGTFTN-----ANVQPSNYSFVAFSADVTP----- 226
DQ 213 KTNADTAATTFIKELKDSAKLT-PGTTAEGQAASQOQPNYSFVGFSDNVTGRTGSQEQ 271
QY 227 --VNYKYARRTVNGDEPS-----SRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTG 281
DQ 272 DLPSNFAKRVVMTTEGLSAGOTQIVSSEPLTDVSWIYSLTGAGSKYTLTFTYGPSTA 331
QY 282 YLYPPYKLVKADANNVGLYKLNNG-NVQOVEFATSTSANNTTANPTPAVDEIKVAV 340
DQ 332 YLYFFIKLVQSDSNKIGLQKLNNSSETLVITFGNETN-----DSGATPAIDDIKVEKT 387
QY 341 LSGLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSNNENNADKI 384
DQ 388 LSNLNFQDNTIEFSVATEE---NKVAPMIGNMYLTSSTNNVDKI 428

RESULT 5
Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Adhesin pmGAL.4 (Fragment).
GN pmGAL.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmGA multigene family of Mycoplasma gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91415.1; -.
DR PROSITE; P500104; EPSP_SYNTHASE_1; 1.
FT NON_TER 584
SQ SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;

Query Match
Best Local Similarity 40.2%; Score 805; DB 2; Length 584;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

QY 11 NPNNGOT-----QLEAARMELTDLINAKAMTSLASQDYAKTEASLSAYSEATVNNLNA 66
DQ 69 NPNNGNTTPEQQLAARLTLDLGTENTNVYALYAKIQSTLTSTAYMTAKTASNTSA 128
QY 67 TLEQLKMAKNLESAINQANTDKTTTDENEHPNLVEAYKALKTTLEQATNLEGLSTAYN 126

```

```

DQ 129 TLDNLSRSTTLQAAIDKAAASNKRTFDSANQPLVYATYNQLKTTLOSKTTSLEGLSENKYS 188
QY 127 QIRNNVLNLYNKASSLITKTLDPLNGGTLDSNEITTANKNNINLTSTINQKTNADALS 186
DQ 189 SIKNLSKLFDAAGSAIAARTLDP-TMGTVPEVMSVTKANEDIMTAVSKLTEWKTNADKF- 246
QY 187 NSFIKKVIQNEQSFVGTFTNAN--VQPSNYSFVAFSADVTP-----VNYKYARRT 235
DQ 247 NDFEKKPL--SKEKLVSTNDRAHQEQPANNSFAGYSVDLTGTSGNSQNLPNMNFQKR 304
QY 236 VVNGD---EPSRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKA 292
DQ 305 VMTSEGOQTGKTALVSSPSATDVSWIYSLAGEGTYLSEYIGPDFAFLYFPYKLVKQ 364
QY 293 ADANNVGLYKLNNGNVQOVEF-----ATST-----S 319
DQ 365 ADSSVALQYSLNKTSSKLINFEPAKTMPTNQDQSENGVATTSTTEGRSSSEVLVADEVA 424
QY 320 ANNTTANPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSNEN 379
DQ 425 AVNEMNPTPTVSDINIAKVTLSGLTGFTEIEFSVPT-----NKVAPMIGNMYLTSNSG 479
QY 380 NADKI 384
DQ 480 SQGKI 484

RESULT 6
Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Adhesin pmGAL.3.
GN pmGAL.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmGA multigene family of Mycoplasma gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -.
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DPE47 CRC64;

Query Match
Best Local Similarity 39.8%; Score 761; DB 2; Length 680;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

QY 9 DANPNNGOTQLEAARME-----LTDLINAKAMTSLASQDYAKTEASLSAYSEATVN 61
DQ 63 DTPNCGGOGMMNATNQELVNAKALSDLLGGSEKTVELYADYAKIRADLTSAYAVAKTS 122
QY 62 NNLNATLEOLKMAKNLESAINQANTDKTTTDENEHPNLVEAYKALKTTLEQATNLEGLS 121
DQ 123 DSSSTLDOQVKTATSTLQTAINTAASDREKFPQONSQLLMAYKVLKDTLKKKEATVMSLN 182
QY 122 STAYNOIRNNVLNLYNKASSLITKTLDPLNGTLLDSNEITTANKNNINLT-----ST 174
DQ 183 QEKYSAILSEINAASSTAEIVKOTLNPVNG-----NLPVVAALNAENTKLEAIKEEK 236
QY 175 INEQKTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARR 234
DQ 237 INSEKSNADLFANYQLYKL---DRYKLMSEGSNNTKQPCNYSFVAFSADIASPNMNFQAR 293
QY 235 TVWNGD-----BPSSRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTGYLYFPYKL 289
DQ 294 TWTADSRWTSPPLPNNLQNSAPLTDVSWIYSLSGTGAKYTLTFTDYGPQTCYLYFPYKL 353

```

```
QY 290 VKAADANNVGLQYKLNNGNQVQVEF-----ATSTSAANTTANPTTAV 331
Db 354 VKTSD--KVGLQYKLNQADPVAQFSEAAATASAPAEATDGRQESAETATANKVNPMPV 411
QY 332 DEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNVAPMIGNIYLSNNENADK 383
Db 412 NTINAKVITLSLNLKFGSNTIEFVPMDDNMKNVAPMIGNIYITSSNDEANK 463

RESULT 7
Q49498 ID Q49498 PRELIMINARY; PRT; 702 AA.
AC Q49498:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PMGAI.3 protein precursor.
GN PMGAI.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62417.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 702 AA; 75537 MW; 2738915FEB57B9F CRC64;

Query Match 37.6%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 4.3e-28;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 11 NPNNGOT----QLEAARMELTDLINAKAMTASLODYAKIEASLSAYSEAEVTNNLNA 66
Db 69 NPNPGNTPEQQLAAARKLTLDLGTENTNVALYADYAKIQSTLSTAYMTAKTASENTSA 128
QY 67 TLEQLKMAKTNLSEAINOANTDKTTDFNEHPNLVEAYKALKTTLEORATNLEGLSTAYN 126
Db 129 TLENLSASTTLOAADKANDKRVDSVNPQLVAAYNNLKTLLKSKTSLGLESENKYG 188
QY 127 QIRNNLDVLYNKASSLITKTLPLNG--GTLDSNEITTANKNINNTLS--TINEQKTA 182
Db 189 GIKHLKSLFDTGSAITAKTLDPTSGRPTLEKVNE---ANNGIKWAISPESLKKWKGA 245
QY 183 DALNSFIKVKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARTVW 237
Db 246 DKF-NEFENKPLSEKLTSTSTAHNQEQPANMSFAAYSVDELTSNQLPNWFAQRYW 304
QY 238 NGD--EPSSRLIANTN-SITDWSWISLAGTNTKYQFSFNSGPGSYGLYFYPKLVKAD 294
Db 305 TSENQPGKTALVSSPVSAITDWSWISLAGEGTKYTLTFEYGPDPNAFLYLPKLVKAD 364
QY 295 ANNVLQYKLNNGNQVQVEF-----ATST-----SAN 321
Db 365 SSSVALQYSLNKTSSKLINFKPAETVSTNQDSENEVATSTTEARSSYKVLVADAAAT'S 424
QY 322 NTANTPTPAVDKIVKIVLSGLRFGONTIELSVPTGEGNMKNVAPMIGNIYLSNNENA 381
Db 425 NNEMHNTPTVSDINIAKVTLSGLTGFENTIEFSVPEG-----KVAPMIGNIYLSNSEQ 479
QY 382 DK1 384
Db 480 VK1 482
```

RESULT 8

```
Q9ZID1 ID Q9ZID1 PRELIMINARY; PRT; 645 AA.
AC Q9ZID1:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE M9 protein.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;

Query Match 37.2%; Score 746; DB 2; Length 645;
Best Local Similarity 42.3%; Pred. No. 9.3e-28;
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;

QY 9 DAPNNGQ-----TOLEAARMELTDLINAKAMTASLODYAKIEASLSAYSEAEVTN 61
Db 44 DYNPGGQGGMMNAASOELAAARMGLTTFVDSKAKNLGLVYDYKKTQDTLTTKAYDAKTVL 103
QY 62 NNLNATLEQLKMAKTNLSEAINOANTDKTTDFNEHPNLVEAYKALKTTLEORATNLEGLS 121
Db 104 DNSSTTQNLNEAKTRLETAIRTAATSKQTFDQHAELVKVYEELKTTLSNETATLAPYA 163
QY 122 STAYNOIRNLDVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTL--STINEQK 179
Db 164 AAQYAGIKRHLGSLYDAGKAITTKLEPVEGDP-LTADVVMANTKIVEAIKDEVLPQK 222
QY 180 TNADALNSFIKVKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTP-----NYKY 231
Db 223 ENATKLADSFVKQVLYKEITGVVEAHN-KAQPANYSFVGISVDITGTVTGQTSIPNWDY 281
QY 232 ARTVW-NGDEPSSRLANT-----NSITDWSWISLAGTNTKYQFSFNSGPGSY 282
Db 282 AQRTIETNGDEP--RSISNTPADQGMVQPLSNVSWISYSLAGTGAKYTLFTYVGPSTGY 339
QY 283 LFYPYKLVKAAADANNVGLQYKLNNGNQVQVEATSTSA-----NNTTANPTPAVDKIVA 337
Db 340 LFYPYKLVWTSQMKLGLEYKLD-----ATEPSAITFGSEQTMNGKTPTVNDINVA 391
QY 338 KIVLSGLRFGONTIELSVPTGEGNMKNVAPMIGNIYLSNNENADKI 384
Db 392 KVTLANLKFGSKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 433

RESULT 9
Q9ZHR9 ID Q9ZHR9 PRELIMINARY; PRT; 644 AA.
AC Q9ZHR9:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE Hypothetical 69.9 kDa protein.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
```


Best Local Similarity 42.98; Pred. No. 1.2e-27;		
Matches 172; Conservative	62; Mismatches 127; Indels 40; Gaps 11;	
QY	11 NPNNQ-----TGLEAARWELTDLINAKAMTILASLDYAKIEASLSAYSEAFYNNN	63
Db	51 NPGDGGMMNAAQELAAARGLTTFVDSKAKNLGLVDYKKTQNTLTAKYDAAKTVDLN	110
QY	64 LNAITLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTTLEORATNLEGLSST	123
Db	111 SSSTQNLEAKTRLETAIRTAATSKOTFDEQHAELVKVYKELKTTLSNETATLAPYADA	170
QY	124 AYQIRNRLVDLYNKASSLTKLIDPLUNGTTLLDSNEITTKANKINNLT--STINEOKTN	181
Db	171 QYAGIKMHLGSLYDAGRAITTKTLEPYEGDP-LTAGAVTMANTKRIVEAIKDEVLNPKKEN	229
QY	182 ADALNSFIKKVIQNNPQSFGVTFNANVPSPNYSFVAFSADTTPV-----NYKVAR	233
Db	230 ATKLAUSFVKQVLVKEKITGVEEAHN-KAQPNATSFVGYSDVITGTANGQTSIPNNWTAQ	288
QY	234 RTVW-NGDEPSSRILANT-----NSITDVSMIYSLAGTNTKTYQFSNYPSTGYLY	284
Db	289 RTIFTNGDEP-RSVSNTPVDGQDMAQPLSNVSIYSLAGTGAKYLEFTYGYSTGYLY	346
QY	285 FPKLYKAAADANVGLQYKLNNGVQOVERTATSTSANNTTAN-PTPAVDEIKVAKIVLSG	343
Db	347 FPKLVNTSQMKGLEYKLDA----TEPSAITFGNEQTMNGRTPTVNDINAKVTILAN	402
QY	344 LRFGQNTIELSVPTGEGNMKNVAPMIGNIYVLSNENNADKI	384
Db	403 LIFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI	438
RESULT 11		
Q49497	PRELIMINARY; PRT; 649 AA.	
AC	Q49497	
DT	01-NOV-1996 (TREMBlrel. 01, Created)	
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)	
DT	01-OCT-2002 (TREMBlrel. 22, Last annotation update)	
DE	PMGAL.2 protein precursor.	
GN	PMGAL.2.	
OS	Mycoplasma gallisepticum.	
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.	
OX	NCBI_TaxID=2096;	
RP	[1]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN-S6;	
RK	MEDLINE=95010739; PubMed=7925999;	
RA	Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,	
RA	Browning G.F., Whithear K.G., Walker I.D.;	
RT	"The organisation of the multigene family which encodes the major cell	
RT	surface protein, PMGA, of Mycoplasma gallisepticum.;"	
RL	FEBS Lett. 352:347-352(1994).	
DR	EMBL; L28424; AAA62416.1; -.	
KW	Signal.	
FT	SIGNAL	
SEQ	SEQUENCE 1 25 POTENTIAL.	
	SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;	
Query Match	37.1%; Score 743; DB 2; Length 649;	
Best Local Similarity 42.3%; Pred. No. 1.3e-27;		
Matches 172; Conservative	62; Mismatches 125; Indels 48; Gaps 11;	
QY	9 DANPNNGQ-----TGLEAARWELTDLINAKAMTILASLDYAKIEASLSAYSEAFYNN	61
Db	48 DTPNGDGGMMNAASQELAAARGLTTFVDSKAKNLGLVDYKKTQNTLTAKYDAAKTVDLN	107
QY	62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTTLEORATNLEGLS	121
Db	108 DNSSSTQNLEAKTRLETAIRTAATSKOTFDEQHAELVKVYKELKTTLSNETATLAPYA	167
QY	122 STAYNQIRNRLVDLYNKASSLTKLIDPLUNGTTLLDSNEITTKANKINNLT--STINEQK	179
Db	168 DAOYAGIKMHLGSLYDAGRAITTKTLEPYEGDP-LTASAVMANTKRIVEAIKDEVLNPK	226

```
QY 180 TNADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTVP-----NYKY 231
Db 227 ENATKLADSFVKQVLYKKEITGVEEAHN-KAOPANYSFVGYSVDITGTTTGTQTSIPNWDY 285
QY 232 ARTVW-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGY 282
Db 286 AORTIFNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTNGAKYTFLETTYGSTGY 343
QY 283 LYPPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSA-----NNTTANPTPAVDRIKVA 337
Db 344 LYPPYKLVNTSDQVGLGLEKLDN-----ATKPSAITFGSDQTMNGKTPVNDINVA 395
QY 338 KIVLSGLRGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADKI 384
Db 396 KVTLANLFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437

RESULT 12
Q8RLX9 PRELIMINARY; PRT; 486 AA.
AC Q8RLX9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hemagglutinin (Fragment).
GN VLHA.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ts-11;
RA Markham P.F.; Kanci A., Czifra G., Sundquist B., Haines P.,
  Browning G.F.;
RT "Molp homolog in Mycoplasma gallisepticum is not essential in vitro.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY065985; AAL58980.1; -.
FT NON_TER 486
SQ SEQUENCE 486 AA; 51844 MW; 94912DD7A09FB911 CRC64;

Query Match 36.9%; Score 739.5; DB 2; Length 486;
Best Local Similarity 38.9%; Pred. No. 1.4e-27;
Matches 171; Conservative 67; Mismatches 123; Indels 79; Gaps 11;

QY 9 DAPNNG-----QTOLEARMELTDLINAKAMTILASLDQYAKIESLSAYSEATVNN 62
Db 62 DTPNGRGGMNAVQGLAAAKTALTLLNGQTEKVGLYNDYAKIKDDLVKAYIAKEISD 121
QY 63 NLNATLEOLKMAKTNLESAINOANTDKTTFDNEHPNLVAYKALKTTLEQRATNLEGLS 122
Db 122 KSHATLOEVNNAKTRLETAIKDANSKTSFGKNPELIKAYDAKQTIYSEMSLNQMD 181
QY 123 TAYNQIRNLDVLYNKASSLITKTLPLNGGTLDSNE---ITTANKNINNTLSTINEOK 179
Db 182 ANFETIKHISNLYKGGDIATLDPPTG---DGPQAVYNOTNEAIVNATSKEDWK 237
QY 180 TNADALNSFIKKVION-----NEQSFVGTETNANVQPSNYSFVAFSADV-----TP 226
Db 238 TNATNLATRFVKOTLNNANLVNE-----TNNQPOPGSYSFVAYSVDLNTGVSTASNT 290
QY 227 VNYKARRTVW---NGDEP---SSRLANTNSITDVSWIYSLAGTNTKYQFSNYPST 280
Db 291 -NWNLAQRKVVWGGSGRTSPSSDANSNPAITDVSWIYNSGANSKYTLTFNMYGPST 349
QY 281 GYLFPYKLVKAADANNVGLQYKLNNGNVQVFEA-----NYKY 315
Db 350 GYLFPYKLVKSDDAQNVGLQYTLNNKPAQRIEAPASPSSGGTAHASPQSPRAAAT 409
QY 316 -----TSTSANNTANPTPAVDRIKVAIKVLSGLRGONTIELSVPTGEGNMKN 364
Db 410 TDVSDSAEGSQATDMSSSSMKKTPTVSDINVASVTLSDLNFGNTIEFSVPMGD---SM 466
```

```
QY 365 VAPMIGNIYLSNENNADKI 384
Db 467 VAPMIGNMYTISNPLAVNQI 486

RESULT 13
Q49468 PRELIMINARY; PRT; 647 AA.
ID Q49468 Q53303;
AC Q49468; Q53303;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Hemagglutinin homolog precursor.
GN PMGA1.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93162830; PubMed=8432610;
RA Markham P.F.; Glew M.D., Whithier K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmGA, a
  hemagglutinin of Mycoplasma gallisepticum.";
RL Infect. Immun. 61:903-909(1993).
DR EMBL: M83178; AAA02996.1; -.
DR EMBL: S55216; AAB25397.2; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match 35.6%; Score 713; DB 2; Length 647;
Best Local Similarity 41.1%; Pred. No. 3.3e-26;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

QY 9 DAPNNGQ-----TQLEARMELTDLINAKAMTILASLDQYAKIESLSAYSEATVNN 61
Db 48 DTPNGGQGMNAASQELAAARMGLTIFDSKARNLGLYVDYKKTQNTLTAKYDAKTVL 107
QY 62 NLNATLEOLKMAKTNLESAINOANTDKTTFDNEHPNLVAYKALKTTLEQRATNLEGLS 121
Db 108 DNSSSTQNLNEAKTRLETAIRTAATSKOTFFDQHAELVKYKELATLTSNETATLAPYA 167
QY 122 STAYNQIRNLDVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTL--STINEOK 179
Db 168 DAQYAGIKMHLGLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLPQK 226
QY 180 TN----ADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTVP----- 227
Db 227 ENATKLADSLSSIVKKTGVEE-----AHNKAQPNYSFVGKRWITTELLDQVFP 279
QY 228 NYKARRTVW-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYPST 278
Db 280 NWDYAQRTIFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTNGAKYTFLETTYGP 337
QY 279 STGYLFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSA-----NNTTANPTPAVD 333
Db 338 STGYLFPYKLVNTSDQVGLGLEKLDN-----ATKPSAITFGSDQTMNGKTPVND 389
QY 334 IKVAKIVLSGLRGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADKI 384
Db 390 INVAKIVTLANLFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 435

RESULT 14
Q9KH15 PRELIMINARY; PRT; 656 AA.
ID Q9KH15
AC Q9KH15;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Adhesin pmGA1.2.
GN PMGA1.2.
```

OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.

RA SHEN Q.C., Bi D.R., Weng C.J.;
RC STRAIN=HS;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
RT gallisepticum strain HS";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275131.1; AAF91413.1; -;
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;

Query Match 35.1%; Score 704; DB 2; Length 656;
Best Local Similarity 41.0%; Pred. No. 9e-26;
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

QY 9 DAMPNNGQT-----OLEAARMELTDLINAKMTLASLDQYAKIEASLSAYSEATVN 61
DB 63 DINPGGGQNMDSAAQELFAARTALTSLLASKNANVENYSDYAKIONTLIAAYTTAEQTS 122
QY 62 NNINATLEQKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQKATNLEGLS 121
DB 123 QNSATLEQKNAKTSALQTAINTANSKQKFDQDHSNLLSYKNLMATLAKKETAVMTLK 182
QY 122 STAYNOIRNLDLYNKASLITKTLDPNGTLLDSNEITTANKNINNTLS--TINEOK 179
DB 183 DPKYSAILDOINGVSKGELVQHTLDPVS-GIVPAANTITEITKIEEIVISEKTLDQK 241
QY 180 TNADALSN--SFT--KKVTQNNQSFVGTFTTANQVSNYSFVAFSADVTVP----- 227
DB 242 NNADQFANYQSFTLDKTKLENVEDA-----KMGQPNYSFVGSVDVTGTSQGTTP 295
QY 228 NYKARTVWNGDEPSSRILANTNS-----ITDVSWIYSLAGNTKYQFSFSGPS 279
DB 296 NMFQRAIFTSQNPQTKVATTGDSQSTAKPLSDVSWIYSLAGTCAKYLEFTYGPS 355
QY 280 TGYLYPYPYKLVKAADANNQYKLVN-NGNQVQVEFATSTANNTTANTPAIDEIKVAK 338
DB 356 TGWLYPYPYKLVKAND--DVGLQYKLSNETLPTIIEGEGT---TTNGPAAVTENINVAK 409
QY 339 IVLGSLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENNAK 383
DB 410 VRLTGLAFGKNTIEFSVP-----MSKVAPMIGNMYITSSDTEINK 449

RESULT 15

Q05122 PRELIMINARY; PRT; 703 AA.
AC Q05122;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Haemagglutinin.
GN PMGAI.9
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whitham K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum";
RL FEBS Lett. 352:347-352(1994).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; U90714; AAB50154.1; -;

SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CBC5 CRC64;
Query Match 34.5%; Score 690.5; DB 2; Length 703;
Best Local Similarity 35.0%; Pred. No. 4.2e-25;
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;

QY 2 CMSITKKDANP-----NNGQT-----OLEAARMELTDLINAK 33
DB 26 CTSATIPTLNPPTNPEKPDPPNPPSGMGNGGNTNPGMDTAQELASAKAALTTLTNRE 85
QY 34 AMTLASLODYAKIEASLSAYSEATVNNNLNATLEQLKMAKTNLESAINOANTDKTTFD 93
DB 86 SEKVGLYVDYAKIKADLTSAITVAKTSDSSSTLVQVKTATSKLOTADDKAASDKQKE 145
QY 94 NEHPNLVEAYKALKTTLEQR-ATNLEGLSSTAYNQIRNLDLYNKASLITKTLDPNG 152
DB 146 QDHKDLLMPYSELKTTLSQKNATVL--LNQPKYSAILNKINSIYAQGEVWIRTLDPSV- 202
QY 153 GTLLDSNEITTANKNINNTLS--TINEOKTNADALSNFSIKKVIQNHQESFVGTFTTANV 210
DB 203 GAIPTAASITTKVNDENKAISENQKPKKDNADAFANYOFFKL---DKTKIMGSMSTNMK 259
QY 211 QPSNTSFVAFSADVTVP-----NYKARTVWNGDEPSSRILANTNS----- 252
DB 260 QPNTSFVGSVGVTCMQSGQTTPNNFNAQRIWSSGAPRAPLASOTETPOAETPMSA 319
QY 253 -----ITDVSWIYSLA 263
DB 320 PQGVEPAQQGDSPPKQASSETQEVSPTPAAEQAAQADTEQPATSQGTPLTDVSWIYSL 379
QY 264 GTNTKYQFSFSGPSNYGPSTGYLYPYPYKLVKAADANNQYKLVNNGVQVQVEFATSTANNT 323
DB 380 GTDVKITFTFTNFGPSNAYLYIPYKLVKSD--SVGLQYKLSNNNNPVALNFGSETNAN-- 435
QY 324 TAMPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENN 380
DB 436 --GPAASVDNINVAKNVLANLNFNGENTIEFSVP-----MNKVAPMIGNMYITSDVAN 485

Search completed: October 8, 2003, 17:31:46
Job time : 100 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:09:46 ; Search time 22 Seconds
(without alignments)

842.206 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSITKDKANPNNGQTGLE.....SSNENNADKIPGYRRPGTFL 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	7.7	682	1	NISP_LACLA
2	143	7.1	857	1	AR56_CANAL
3	142	7.1	2660	1	YEEJ_ECO57
4	138.5	6.9	1314	1	SWTJ_YEAST
5	137	6.8	719	1	YMA1_YEAST
6	135.5	6.8	1085	1	Y441_YEAST
7	132	6.6	1271	1	Y338_MYCGE
8	131.5	6.6	1164	1	BAG_STRAG
9	129	6.4	1251	1	RBP2_PLAVB
10	128.5	6.4	2334	1	WPA_BACSU
11	128.5	6.4	3712	1	LMA_DROME
12	128	6.4	995	1	YI09_YEAST
13	127	6.3	544	1	FLGK_BUCAP
14	127	6.3	956	1	YEF3_YEAST
15	126.5	6.3	444	1	PST1_YEAST
16	126.5	6.3	1630	1	MSF1_PLAFK
17	126.5	6.3	1639	1	MSF1_PLAFW
18	125.5	6.3	1858	1	P3K2_DICDI
19	124.5	6.2	821	1	LN5_CAEEL
20	124.5	6.2	1010	1	YK1_CAEEL
21	124.5	6.2	1018	1	FBA_STRAU
22	124	6.2	749	1	MAD1_YEAST
23	124	6.2	750	1	YKS7_YEAST
24	124	6.2	823	1	NSP1_YEAST
25	124	6.2	866	1	MYSP_SCHJA
26	124	6.2	1117	1	YV96_YEAST
27	124	6.2	1159	1	N124_SCHPO
28	124	6.2	1433	1	CAT8_YEAST
29	123.5	6.2	537	1	ARP_PLAFA
30	123.5	6.2	1902	1	P2P_LACLC
31	123	6.1	1260	1	ALS1_CANAL
32	123	6.1	1637	1	MRSP_STRAU
33	123	6.1	1790	1	USOL_YEAST

ALIGNMENTS

RESULT 1

ID	NISP_LACLA	STANDARD;	PRT;	682 AA.
AC	Q07596;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Nisin leader peptide processing serine protease nisp precursor (EC 3.4.21.-)			
GN	NISP.			
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.			
OX	NCBI_TaxID=1360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NIZO R5;			
RX	MEDLINE=93239683; PubMed=8478324;			
RA	van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J., Kuipers O.P., de Vos W.M.;			
RA	*Characterization of the Lactococcus lactis nisin A operon genes nisp, encoding a subtilisin-like serine protease involved in precursor processing, and nispR, encoding a regulatory protein involved in nisin biosynthesis.;			
RT	J. Bacteriol. 175:2578-2588(1993).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=6F3;			
RX	MEDLINE=94213458; PubMed=8161176;			
RA	Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K., Hammelmann M., Entian K.-D.;			
RA	*Regulation of nisin biosynthesis and immunity in Lactococcus lactis 6F3.;			
RT	Appl. Environ. Microbiol. 60:814-825(1994).			
RL	[3]			
RP	3D-STRUCTURE MODELING.			
RX	MEDLINE=95357326; PubMed=7630881;			
RA	Siezen R.J., Rollemans H.S., Kuipers O.P., de Vos W.M.;			
RA	*Homology modelling of the Lactococcus lactis leader peptidase Nisp and its interaction with the precursor of the lantibiotic nisin.;			
RT	Protein Eng. 8:117-125(1995).			
CC	- FUNCTION: CLAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.			
CC	- PATHWAY: Nisin biosynthesis; last step.			
CC	- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (potential).			
CC	- SIMILARITY: Belongs to peptidase family S8.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L11061; AAA25200.1; -			
DR	EMBL; X76884; CAA54210.1; -			

34	122.5	6.1	507	1	FLIC_SALON	Q06974 salmonella
35	122	6.1	650	1	PDR4_YEAST	P19880 saccharomyc
36	122	6.1	1672	1	PMPB_CHLMU	Q9pjy2 chlamydia m
37	121.5	6.1	425	1	YBY0_YEAST	P38272 saccharomyc
38	121.5	6.1	1140	1	YV96_YEAST	Q04893 saccharomyc
39	121	6.0	491	1	TOLC_SALEN	Q54001 salmonella
40	121	6.0	901	1	PIP_LACLA	P49022 lactococcus
41	121	6.0	979	1	P115_MYCHR	P41508 mycoplasma
42	121	6.0	1111	1	KIPL_YEAST	P28742 saccharomyc
43	121	6.0	1902	1	P2P_LACPA	Q02470 lactobacill
44	120.5	6.0	776	1	VP4_ROTCH	Q02945 equine rota
45	120.5	6.0	800	1	INLA_LISMO	P25146 listeria mo

RP PIR; S44131; S44131.
RC HSP; P29600; IGCI.
RX MEROPS; S08.059; -.
RA InterPro; IPR006192; LPXTG.
RA InterPro; IPR000209; Peptidase_S8.
RF Pfam; PF00082; Peptidase_S8; 1.
RT PRINTS; PR00723; SUBTILISIN.
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate - N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate - ADP + N-acetyl-L-
CC glutamate 5-phosphate
CC -1- PATHWAY: Arginine biosynthesis; second step.
CC -1- PATHWAY: Arginine biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X98880; CAA67383.1; -.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; Acglut_kinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinas; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdh_dh; 1.
DR Pfam; PF02774; Semialdh_dh; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARG; 1.
DR Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
KW Mitochondrion; Multifunctional enzyme; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT REDUCTASE
FT ACT_SITE 669 669 BY SIMILARITY
SQ SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C4C71 CRC64;

Query Match 7.1%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.72;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;

QY 46 IEASLSAY-----SEAEVNN---NLNATL-----EOLKMAKTNLESAINQAND 88
DB 208 VEAAINSGLPILTSIAETSSGGLNVDVAAAGELAREFEPLKIVYLNKGGIINGNTG 267
QY 89 K-----TTFDNEHPLV-EAYKALKTTLEOR-----ATNLEGLSSTAYNQIRNNLDVLYNK 138
DB 268 EKVSAINLDEEYEDLLKESVKGVTGKIKIHDHLLQHLPRSSVA-----IIDVNDL 320
QY 139 ASSLIITKLDPLNGTLLDSNEITANKNIN-NTLSTINQKTNADALSNSFTKVKIQNN 197
DB 321 QKELFTDS---GAGTL-----IRRGYRLINRSLRDFG---NPDLLRALLR----- 361
QY 198 EQSFVGTFTNANVOPSNYSFVAFADYTPVNYKYARTVWNGDPSRILANT----- 250
DB 362 -----DPEIKTGKVSASYLKFLDSVQFKS-----YGDPE-LEVLAIIVEQNDKI 405
QY 251 -----NSITD-----VSWIYSLAGTNTKYQSFNSYG-PSFGYL 283
DB 406 PKLDFEFLSKTGWLNNTDINFNALIKDDYSQLCVWVNDANLFWYFSKDGSPAKNGQI 465

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 64385 / 1001;
RX MEDLINE-97195775; PubMed-9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5, 6 gene of
RL Candida albicans";
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate - N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate - ADP + N-acetyl-L-
CC glutamate 5-phosphate
CC -1- PATHWAY: Arginine biosynthesis; second step.
CC -1- PATHWAY: Arginine biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X98880; CAA67383.1; -.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; Acglut_kinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinas; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdh_dh; 1.
DR Pfam; PF02774; Semialdh_dh; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARG; 1.
DR Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
KW Mitochondrion; Multifunctional enzyme; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT REDUCTASE
FT ACT_SITE 669 669 BY SIMILARITY
SQ SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C4C71 CRC64;

Query Match 7.1%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.72;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;

QY 46 IEASLSAY-----SEAEVNN---NLNATL-----EOLKMAKTNLESAINQAND 88
DB 208 VEAAINSGLPILTSIAETSSGGLNVDVAAAGELAREFEPLKIVYLNKGGIINGNTG 267
QY 89 K-----TTFDNEHPLV-EAYKALKTTLEOR-----ATNLEGLSSTAYNQIRNNLDVLYNK 138
DB 268 EKVSAINLDEEYEDLLKESVKGVTGKIKIHDHLLQHLPRSSVA-----IIDVNDL 320
QY 139 ASSLIITKLDPLNGTLLDSNEITANKNIN-NTLSTINQKTNADALSNSFTKVKIQNN 197
DB 321 QKELFTDS---GAGTL-----IRRGYRLINRSLRDFG---NPDLLRALLR----- 361
QY 198 EQSFVGTFTNANVOPSNYSFVAFADYTPVNYKYARTVWNGDPSRILANT----- 250
DB 362 -----DPEIKTGKVSASYLKFLDSVQFKS-----YGDPE-LEVLAIIVEQNDKI 405
QY 251 -----NSITD-----VSWIYSLAGTNTKYQSFNSYG-PSFGYL 283
DB 406 PKLDFEFLSKTGWLNNTDINFNALIKDDYSQLCVWVNDANLFWYFSKDGSPAKNGQI 465

PTR; S44131; S44131.
DR HSP; P29600; IGCI.
DR MEROPS; S08.059; -.
DR InterPro; IPR006192; LPXTG.
DR InterPro; IPR000209; Peptidase_S8.
RF Pfam; PF00082; Peptidase_S8; 1.
RT PRINTS; PR00723; SUBTILISIN.
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate - N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate - ADP + N-acetyl-L-
CC glutamate 5-phosphate
CC -1- PATHWAY: Arginine biosynthesis; second step.
CC -1- PATHWAY: Arginine biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X98880; CAA67383.1; -.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; Acglut_kinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinas; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdh_dh; 1.
DR Pfam; PF02774; Semialdh_dh; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARG; 1.
DR Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
KW Mitochondrion; Multifunctional enzyme; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT REDUCTASE
FT ACT_SITE 669 669 BY SIMILARITY
SQ SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C4C71 CRC64;

Query Match 7.7%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.13;
Matches 75; Conservative 61; Mismatches 130; Indels 91; Gaps 13;

QY 58 ETVNNINLNLATLEQLKMAKTNLESAINQANDTKTTFDNEHPLV-EAYKALKTTLEQRTNL 117
DB 38 ELINHSNAILSSSTEGSTDSINLGAQSPAVKSTRTT-----LDVTGAATLLQTSVAQK 93
QY 118 EGLSSTAYNQIRNNLDVLYNKASSLIITKLDPLNGTLLDSNEITANKNINNTLSTINE 177
DB 94 EMKVSLOEQTSSE-----FSRDSVTNKEAVPVSKEDELQSEVVVSTSIQKN-KILDN 148
QY 178 QKTADALSNSFIKVIQNNQSQSVGFTTNANVOPSNYSFVAFADYTPVNYKYARTV- 236
DB 149 KKRRANFTVSSPLIKERPSNKSADGVINDS-----ASPLSYRKAKEWS 193
QY 237 -----WNGDEPSRILANTNSITDYSWYISLAGTNTKYQSFNSYGPS 279
DB 194 LRQPLKNQVZEAQPLLLISNSSEKASVYTNHDFWYQW-----DMKY---VTNNGES 243
QY 280 TGYLYFPYKLVKAADANNVGLQYKLINGNVQOQVEFATISAN-----NT 323
DB 244 YA-LXQPSKKI-----SVGI---IDSGIMEEHPDLSNLSGNYFKNLVPGKGFDENEPDE 293
QY 324 TANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIG-NIYLSNEN 379
DB 294 TGNPSDIVD-----KMGHGTGVAQGITANGNIGLVAFGITVNIYRVFGEN 338

RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ARG5.6 protein, mitochondrial precursor (Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-L-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)).
GN ARG5.6.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]

FT	DOMAIN	738	834	BIG-1 1.	
FT	DOMAIN	840	929	BIG-1 2.	
FT	DOMAIN	931	1033	BIG-1 3.	
FT	DOMAIN	1042	1132	BIG-1 4.	
FT	DOMAIN	1134	1236	BIG-1 5.	
FT	DOMAIN	1245	1335	BIG-1 6.	
FT	DOMAIN	1337	1439	BIG-1 7.	
FT	DOMAIN	1448	1539	BIG-1 8.	
FT	DOMAIN	1548	1652	BIG-1 9.	
FT	DOMAIN	1653	1750	BIG-1 10.	
FT	DOMAIN	1751	1855	BIG-1 11.	
FT	DOMAIN	1856	1957	BIG-1 12.	
FT	DOMAIN	1963	2056	BIG-1 13.	
FT	DOMAIN	2065	2156	BIG-1 14.	
FT	DOMAIN	2157	2252	BIG-1 15.	
FT	DOMAIN	2254	2355	BIG-1 16.	
SQ	SEQUENCE	2660 AA;	280062 MW;	01EB92A08F5C09D2 CRC64;	

Query Match 7.1%; Score 142; DB 1; Length 2660;
 Best Local Similarity 23.7%; Pred. No. 3.3;
 Matches 99; Conservative 60; Mismatches 175; Indels 84; Gaps 18;

QY	3	MSITKKDA--NPNGQT-----QLEAARMEIIDLINAKAMTILASLDYAKIEASLSAY	54
DB	1162	LTAIVKDPSPHPVAGITVNTMPDVAANFTLEN--NGIAITQANGEARVTLGKKAGTH	1219
QY	55	SEAEVTNNNLNATLEQLKMAKTNLESAINQANT-----DKTT-----PONEH	96
DB	1220	TVTATLGNN-NASDAQPVTVADKDSAVVVLQTSKAEIIGNGVDETTLTATVRDPFN--	1276
QY	97	PNLVEAYKALKTTLEQRATNLEGLSSPAYNQIRNNLDVLYNKASSLTKTKTLDLPUNG	156
DB	1277	-----AVKDLQVTF--STN---PADTQLSQSKSNTND-----SGVAEVTFK-----GTVL	1316
QY	157	--DSNETTANKNINWTLSTINQKTNADALSNSFKVKVIONNEQSPVGTFTNANVOPSN	214
DB	1317	GVHTAEATLPGNGNDTKIVNIAPDASNAQVTLNPAQVYVNNNSDSVQLTATVRD--PSN	1374
QY	215	YSFVAFSAADTPVNYKYARTVWNGDPSSRIILANTNSITDVSWIYSLAGTNTKYQSFES	274
DB	1375	HPVAGITVNTMPDVAANFTL-----ENNGIATQANGEARVTLGKKAGTHT-VTATLS	1429
QY	275	NYGPSTGYLYFPKLYKAADANNVGLQYKLNNGNVQVPEAT-----	316
DB	1430	NNNTSDSQ---PVTVEADKTSALVWLQISKNEITNGVDSATLTATVRDQFDNEVNNLPV	1486
QY	317	--STSANNTTANTTPA-VDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKVPIMGN	371
DB	1487	TFTASSGLTLTPGESNTNESGIAQTLAGVAFGEQTVTASLANNGASDNKTVHFIGD	1544

RESULT 4
 SWIL_YEAST STANDARD; PRT; 1314 AA.

ID	SWIL_YEAST	STANDARD;	PRT;	1314 AA.
AC	P09547;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Transcription regulatory protein SWI1 (SWI/SNF complex component SWI1)			
GN	(Transcription regulatory protein ADNR6) (Regulatory protein GAM3).			
DE	ADNR6 OR SWI1 OR GAM3 OR YPL016W OR LPAR1.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RP	MEDLINE=89057455; PubMed=3143101.			
RX	O'Hara P.J., Horowitz H., Eichinger H., Young E.T.;			
RA	"the yeast ADNR6 gene encodes homopolymeric amino acid sequences and a			
RT	potential metal-binding domain.;			
RL	Nucleic Acids Res. 16:10153-10170(1988).			
RN	[2]			

NUCLEIC ACIDS RES. 10:10153-10170(1982).
[2]
RN

SEQUENCE FROM N.A.
 RC STRAIN-S288c / AB972;
 RX MEDLINE-97313271; PubMed-9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoerge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., Dipaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marache R., Messinguy F., Mewes H.-W., Mirtipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Newtich U., Oefner P., Pearson D.,
 RA Patel F.X., Pohl T.M., Purnelle D., Schafer M., Scharif M.,
 RA Scherens B., Schramm S., Schroeder M., Sidcu A.M., Tettelin H.,
 RA Ureastazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wamburt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI.";
 RL Nature 387:103-105(1997).
 [3]
 RN CHARACTERIZATION.
 RP MEDLINE-92154671; PubMed-1339306;
 RX Peterson C.L., Herskowitz I.;
 RA "Characterization of the yeast *SWI1*, *SWI2*, and *SWI3* genes, which
 RT encode a global activator of transcription.";
 RL Cell 68:573-583(1992).
 CC -1- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE *SWI1*/*SNF*
 CC COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER
 CC OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE
 CC BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.
 CC -1- SUBUNIT: COMPONENT OF THE *SWI1*/*SNF* GLOBAL TRANSCRIPTION ACTIVATOR
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 ARID domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X12493; CAA31013.1; -
 DR EMBL: U33335; AAB68089.1; -
 DR PIR: S05728; TNBYR6.
 DR PDB: 1KK5; 04-DEC-02.
 DR TRANSFAC: T01279; -
 DR SGD: S0005937; SWI1.
 DR InterPro: IPR001606; ARID.
 DR Pfam: PF01388; ARID.1.
 DR SMART: SM00501; BRIGHT; 1.
 DR Transcription regulation: Activator; DNA-binding; Nuclear protein;
 KW Zinc-finger; 3D-structure.
 FT DOMAIN 5 65 ASN/THR-RICH.
 FT DOMAIN 337 385 GIN-RICH.
 FT ZN_FING 1241 1258 C4-TYPE.
 FT ZN_FING 1241 1258 C4-TYPE.
 SQ SEQUENCE 1314 AA; 147938 MW; F442D5A82013CDBD CRC64;
 Query Match 6.9%; Score 138.5; DB 1; Length 1314;
 Best Local Similarity 20.4%; Pred. No. 2.2; Indels 91; Gaps 15;
 Matches 70; Conservative 56; Mismatches 126;
 QY 11 NPNNGTQLEAARMELTDLINAKMTLASLDQYAKIESLSAYSEATVNNLNATLEQ 70
 DB 57 NNNNNNTGASGV-----DFQNFDPKPFQDQNLDSNNNSNNNNNNNS--N 104
 QY 71 LKMAKTNLES---AINQANTDKT-----FDNEHPNLVEYKALKTTLEORATNLSGL 120
 DB 105 TVASSTNFTSPTAVVNAAPANTVGKAAFNQONSPQNSPYDSNNNSN-----TNLSL 159

QY 121 SSTAYNOIRNNLYDKASSLITKTLDPLNGTLLDSNEITTANKNNINNTLSTINEQKT 180
 DB 160 SPQAI-LAKNSIID-----SNLPLQAOQLYGG-----NNNNSTGIANDNVI 202
 QY 181 NADALSNSFKKVIQNNNEQSGVGTFTNANVOP-SNYSFVAFSADVTVPVKKYARRVWNC 239
 DB 203 TPHEITN--VQISQNSSSSTPN--TNSNSTPNAQOFLPFNNSASN-----NG 247
 QY 240 DEPSRLTANTSDTSDYWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKKAADANNVG 299
 DB 248 NLTSNQLISN-----YASNSMDRSSASNEFVPT-----SDNNNS 285
 QY 300 LQYKL-----NNGNVQOQVEPATSTSANNTTANPTPAVDE 333
 DB 286 NNHNRRNSNNKTSNNNNVTAPPAATPANTNNSTNSANTVFESE 328
 RESULT 5
 YM41_YEAST STANDARD; PRT; 719 AA.
 ID YM41_YEAST
 AC Q03213;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 79.4 kDa protein in ALD2-DDR48 intergenic region.
 GN YMR172W OR YMR010.02.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288c / AB972;
 RX PubMed-9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Tagela K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
 RL Nature 387:90-93(1997).
 CC -1- SIMILARITY: LOW, TO YEAST MSN1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z49808; CAA89905.1; -
 DR PIR: S55119; S55119.
 DR TRANSFAC: T04601; -
 DR SGD: S0004783; HOT1.
 DR GO: GO:0000228; C:nuclear chromosome; IDA.
 DR GO: GO:0006972; P:hyperosmotic response; IGI.
 KW Hypothetical protein.
 SQ SEQUENCE 719 AA; 79415 MW; 4652BE93743D5A54 CRC64;
 Query Match 6.8%; Score 137; DB 1; Length 719;
 Best Local Similarity 19.8%; Pred. No. 1.2;
 Matches 105; Conservative 72; Mismatches 175; Indels 178; Gaps 23;
 QY 3 MSTTKDANPNNGTQLEAARMELTDL-----INAKMTLAS----- 39
 DB 38 MPTTLKGYRNNQVNEIDAIGINL-DLSLPHISPTTGSSEASGNSASTLRNDGALDGG 96
 QY 40 -LDYAKIEA-----SLSSAYSEATVNNLN-----TLEQKAKTNLES 82
 DB 97 LLRTSAAISATGPTSQPTETIGEKLSNEERVNSVNSASNSTTACTGRLMSQSLTN-DSPS 155

QY 83 NOANTDK-----TTFDNEHPNLV---EAYKALKTTLEQRATNLEGLSSTAY----- 125
 Db 156 NEISTDQLKIFORDMDEMGARMEIEESFNKLSNKTAEQNTWVLUKQDNYKVMKNLNTLL 215
 QY 126 -----NQIRNL-VOLYNKASSLITKTLDPL--NGGTLDSNEIITANKN--- 167
 Db 216 KLVAQPSARPSNTNNAKLAIELLNSISAVSSAYLQKMNNGSGRQHTADLCTGDSNTHS 275
 QY 168 -----INNTLSTNEQTKNA-----DALNSNFIKKVI 194
 Db 276 GINQHRRTNGTIDVNTNTAQLNQFSNALNTILPDQHRNNVNSQINQSLPNROLGPVI 335
 QY 195 -----ONNEQSPV-CTFTNANVQPSNYSFVAFSAD-----VTPVNYKYARRTWNQDEPS- 243
 Db 336 NTOAQNOSQVLIHNTHTNQVNRSPISFPNASTDKPKLNPNGIKRRRRNTQSNNAST 395
 QY 244 -----SRILANTNSITDVSIIYS-----LAGTNTKYQF-SFSNYPGSGY 282
 Db 396 NDHASAAQKPSALSPLNSHNSHTSMYNTSSIHSGVTSASNSPHDLNSLNNFGTTLAL 455
 QY 283 LYFPKYLKADANNVGLQYKLNNGVQOQVEFATSTSNANTANTPTPAVEIKVAKIVLS 342
 Db 456 SLPSLALDNASPPNQNVIPPIINNTQOPLSFSQLINQDSTTS----- 498
 QY 343 GLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSSNENNAKIPGYRRPCT 392
 Db 499 -----EL-LPSGKSGVNT-----NIV-----NRNRASTLPSYPKPMT 529
 RESULT 6
 CUT7_SCHPO
 ID CUT7_SCHPO STANDARD; PRT; 1085 AA.
 AC P24339;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein cut7.
 GN CUT7 OR SPAC25G10.07C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91015362; PubMed=2145514;
 RA Hagan I., Yanagida M.;
 RT "Novel potential mitotic motor protein encoded by the fission yeast
 cut7+ gene.";
 RL Nature 347:563-566(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lechach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
 CC FORM A SHORT SPINDLE THAT ELONGATES TOWARD THE NUCLEUS AT
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X57513; CAA40738.1; -;
 DR EMBL: Z70691; CAA94636.1; -;
 DR PIR: T38378; T38378.
 DR HSP: P17119; 3KAR.
 DR GeneDB: SPombe; SPAC25G10.07c; -;
 DR InterPro: IPR001752; kinesin_motor.
 DR Pfam: PF00225; kinesin; 1.
 DR PRINTS: PRO0380; KINESINHEAVY.
 DR SMART: SM00129; KISC; 1.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN2; 1.
 DR KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).
 FT NP_BIND 897 955 COILED COIL (POTENTIAL).
 FT REPEAT 987 998 ATP (BY SIMILARITY).
 FT REPEAT 999 1010
 FT MOD_RES 1011 1011
 FT PHOSPHORYLATION (BY CDC2) (BY
 FT SIMILARITY).
 FT CONFLICT 34 61 SASNPKRREPPTIDTGYDRSDTNSPT ->
 FT LRAILGNVSLLLTL (IN REF. 1).
 SQ SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;
 Query Match 6.8%; Score 135.5; DB 1; Length 1085;
 Best Local Similarity 20.9%; Pred. No. 2.5;
 Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;
 QY 14 NSQ-TQLEARMELTDLINAKA-MTLASL---QDYAKIEASISLSSAYSEATVNNLNAT 67
 Db 623 NGYTFLLDNFASMEELLNTHSNQLLIISMTKITEHQSLDEALQSARSSCAVPSNLDLI 682
 QY 68 LEQLMAKTNLESAINQANTDKTTFDNEHPN-----LVKAYKALKTT---LEQRAFNLEG 119
 Db 683 VSELKDRKNSLLDALEHSLQDISMSQKLGNGISSELIELQDKMESYQLVQELRSLYN 742
 QY 120 LSTAYNQIRNNVLVDLYNNKASSLITKTLPLNGTGTLDSNEIT-----TA 164
 Db 743 LQTHSEESQKELMYGVNRDIDALVKTCTTSLNDADIILSDYISDQKSPESKQDQLIANI 802
 QY 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVCTFFNANVQPSN 214
 Db 803 GRIVSNLQEQEESLYTKADILHSHLNDNTSNIRKANEIMNKRSEFL---RNA----- 853
 QY 215 YSFVAFSADVTVPVNYKYARRTVWNG-----DEPSSRLANTNSITDVSITYSLA----- 263
 Db 854 ----ASQAEIVGANKRIQKTWENGSQLDLSKSKAIHSNRSMYD-----HCLALAESQKQ 905

QY 264 GINTRYQ-----FSSNYGPGTGYLYFPYKLVKAADANNVGL--QYKLNNGNVQ 310
 DB 906 GVNLEVQTLRLQKVKHSEDNTRKEHQOL---LDLLESVLGNNDNLIDSIKTPHTELO 962
 QY 311 QVE-----FATSSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 351
 DB 963 KTDHVLKGTTSLANHNTNELLGLGDESICNLETTEDTSVLKLETGTPSPKRELPATPS 1022
 QY 352 -----ELSVPTGGNNKVPKMGNIYLLSSNENNADKIPGYRRP 390
 DB 1023 WTRDSSLIKETTNLSDSKKFEVRETYTSSNQTNEPDV--YDKP 1064

RESULT 7
 Y338_MYCGE
 ID Y338_MYCGE STANDARD; PRT; 1271 AA.
 AC P47580;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein MG338 precursor.
 GN MG338.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
 RT "The minimal gene complement of Mycoplasma genitalium."
 RL Science 270:397-403(1995).
 RN [2]
 RC STRAIN-ATCC 33530 / G-37;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing."
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MG307 / MG309 / MG338 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U39715; AAC71563.1;
 EMBL; U01809; AAD12341.1;
 TIGR; MG338;
 PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1;
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 Complete proteome.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1271 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT LIPID 27 27
 SQ SEQUENCE 1271 AA; 142492 MW; FCE6042067310A70 CRC64;

Query Match 6.6%; Score 132; DB 1; Length 1271;
 Best Local Similarity 22.2%; Pred. No. 4.7;
 Matches 84; Conservative 57; Mismatches 140; Indels 98; Gaps 19;

QY 8 KDANPNNGQTOLAEARMELTDLINAKAMTILASLDYAK--TEASLSAYSSEAETVNNNLNA 66
 DB 264 QDPESGGQTN-----SAKL-----FEEDKQLQTANSSGSSQSSAATNTRND 306
 QY 67 TLE-----OLKMAKTNLESANQANTDKTTFDNEHPNLVEAYKALKTKTLQORATNLE 118
 DB 307 LLDPENKYSASKVLISKNNILSVLKTYNLSAAVIDOYH-YLLNNKTELTTTITTTTGTG 365
 QY 119 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGTGLDNEITANK--NINNTLSTIN 176
 DB 366 GTSSNL-----NPLDKFIKSSATTVMK-----SMTKSOEVTSDNNGFNKSEFLKIN 415
 QY 177 E--OKTNADALSN--SFIKKYQIONNEOSFVGTFTNANVPQSNYS-----FVAESAD 223
 DB 416 PSLSSGSDNSNTQSFQKQVQALNNSQTATIFDAVRMESNSQAOVVTSLNLSLSK 475
 QY 224 VTPVNYKYARTVW--NGDE-----PSSR-----ILANTIN 251
 DB 476 TT---OKQOQKPVYVRGDDAIYAFHIDGNGTFLNENSPKKNFEKQAEVLLMRFLQGTN 532
 QY 252 SIT--DVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGYKLNNGNV 309
 DB 533 NFSKDNVSFVDLFGSNESEF--SWARNR--TTLKLYTALTMTLENGTSN-----NNGQK 583
 QY 310 QQVEFATSTSNANTTANPT 328
 DB 584 DVCDLAKLLKNNLTSET 602

RESULT 8

BAG_STRAG

ID BAG_STRAG STANDARD; PRT; 1164 AA.

AC P27951;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE IGA FC receptor precursor (Beta antigen) (B antigen).

GN BAG.

OS Streptococcus agalactiae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1311;

[1]

SEQUENCE FROM N.A., AND SEQUENCE OF 38-48.

RC STRAIN-LA239;

MEDLINE-9132121; PubMed=1857207;

Jeliström P.G., Chhatwal G.S., Timmis K.N.;

"The Iga-binding beta antigen of the c protein complex of Group B

streptococci: sequence determination of its gene and detection of two

blinding regions."

Mol. Microbiol. 5:843-849(1991).

[2]

IDENTIFICATION OF IG-LIKE DOMAIN.

MEDLINE-9703265; PubMed=8880921.

Bateman A., Eddy S.R., Chothia C.;

"Members of the immunoglobulin superfamily in bacteria."

Protein Sci. 5:1939-1942(1996).

CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by

an amide bond (Potential).

CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC

EMBL; X59771; CAA42442.1;

PIR; S15330; FCSOAG

InterPro; IPR004829; Csurface-antigen.

InterPro; IPR005877; Gpos_Y5IRK.

DR

```
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR ProDom; PD153432; Surface_antigen; 1.
DR SMART; SM00409; Ig; 1.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR TIGRfams; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS50835; IG_LIKE; FALSE_NEG.
KW Cell wall; Peptidoglycan-anchor; Receptor; Repeat; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 37
FT CHAIN 38 1135 IGA FC RECEPTOR.
FT PROPEP 1136 1164 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 434 534 IG-LIKE.
FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).
FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
FT DOMAIN 827 945 PRO-RICH REPEATS.
FT SITE 1132 1136 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1135 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1164 AA; 131051 MW; 65DE94AF720A5474 CRC64;

Query Match 6.6%; Score 131.5; DB 1; Length 1164;
Best Local Similarity 22.8%; Pred. No. 4.5;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY 7 KKDANPNNGQOLEA-ARMELTDLINA-KAMTLASLODYAKIEASLSAYSE-----AET 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 RQAQAQDKKEDAEVKVREEELGKFSSTKAGLDQEIQHVKKETSSEENTOKVDEHYANS 289
QY 60 VNNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNL-----VEAYKALKT 108
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 LQNLAKSLEELDRATTN-----EQATQVKNOFLNAQKLKEIQPLIKETNVKLYKAMSE 344
QY 109 TLEQRTNLEGLSTAYNQIRNVLNLYNKASSLITKTYLDPLNGTGLDSNEITTANKNI 168
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 SLEQVEKELKHNSA-----NLBDLVAKSKEIVREYEGKLNQSKNLP--ELKOLEEFA 395
QY 169 NNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQESFGVTGTNANVOPSNYSFVAF 220
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 HSKLKQVDEDFRKKFKTSQVTPKKRVARDLAANNQOKI-----ELTVSPENI----- 445
QY 221 SADVTPVNYKYARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFS-----N 275
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
446 -----TVYEGEDVKFTVTAKSDDSKTTLDPSDLL-----TRYNPSVSDRISTN 487
QY 276 YGPSTG---YLYFPYKLYKAADANNVGLQYKLNNGVVOQVEFATSTSANNTANP-TPAV 331
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
488 YKTNFDNHKIAEITIKNLKLNESQPTVLKAKDDSGNVVEKFTTITVQKKKEKQVKTPEQ 547
QY 332 DEIKVAKIV 340
Db : : : : :
548 KDSKTEERV 556

RESULT 9
RBP2_PLAVB
ID RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
```

```
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites.";
RL Cell 69:1213-1226(1992).
CC !- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC !- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M8098; AAA29744.1; -.
KW Malaria; Receptor; Membrane.
FT NON_TER 1
FT NON_TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.4%; Score 129; DB 1; Length 1251;
Best Local Similarity 21.1%; Pred. No. 6.8; Indels 122; Gaps 20;
Matches 89; Conservative 70; Mismatches 140;

QY 16 QTQLEAARMELTDLINA-KAMTLASIQDYAKIEASLSAYSEAE-----TVNNLNAT 67
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
424 ETNLESVKHNLSEITNIK-----QGGEKI-----YSRAKDIMQIKATSEYTAET 469
QY 68 LEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRTNLEGLSTAYNQ 127
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
470 LEKVADQDSNVYVNLNQITTER-----NLIVTEKRNLANGIDSTITNIEG-----ALKE 517
QY 128 IRNNL-----VDLYNKASSLITKTYLDPLNGTGLDSNEIT--TANKN 167
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 SKGNTYEIGLEKLEIGKRNKRLKVDITRKS-----INSTVG--NFSLFNFDLQVDFNKN 572
QY 168 INNTLSTINEQKTADALSNSF---IKRVIONNEQSFVGTFTNANVOPSNYSFVAFSADV 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
573 INDYENKGE-----IYNEFEGSLNKISEN-----LRNASENTSDYN-----SAKT 613
QY 225 TPVNYKYARRTVWNGDEPSSRILANTNSITDVSWIY-----SLAGTN--TKYQFSFNYG 277
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
614 LRLEAQKREKVNLLNKEEANKYLRDVKVESPFRFENKESLDKINEMIKKQLTVNEGH 673
QY 278 PSTGYLYFPYKLYKAADANNVGLQYKLNNGVVOQVEFATSTSANN-----TTA-- 325
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
674 GNVKOLVENIK--ELVDENNLSDIKQATGRNKEEQKTHSTLKNKAKTILGHVDTSAKY 731
QY 326 -----NPTPAVDE-----IKVAK-----IVLSGLRFQONTIELSVPTGEGNMKV 365
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
732 VGKITPELALTELLGDAKLTAQELKFESKNVNVLETNMSKNTNLDVHKNIQDAYKV 791
QY 366 A 366
Db : : : : :
792 A 792

RESULT 10
WAPA_BACSU
ID WAPA_BACSU STANDARD; PRT; 2334 AA.
AC Q07833;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Wall-associated protein precursor.
GN WAPA OR NI7G.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
```



```

Db 1888 GGDILTESYEDANGNRRTIINSSA---GKVOYEGKL-NQLVKTHEDGVITYEYDGF 1943
QY 151 NG-----GTLLD-SNEITTANKNNINLTSTINEQKTNADALSNSFIKKVIONNEQSFVGTFF 205
Db 1944 GNRKTVTTIKDGSSKTVNASFNIMHQLTKVNDESISYD-----KNGNRSTSDGKF 1992
QY 206 TNANVPQSNYSFVAESADVTP-VNKKYARRIVWNGDEPSSRLANTNSITDVSWIYSLAG 264
Db 1993 TYTWDADNLTAVTKKDGKPFATYKY-----DEKGNRIQRTVNGKV-TNYPYDGD 2043
QY 265 TNTKYQFSFN-----YGPS-----TGVLFPYKLVKAADANNVGLQKLN-NGN 308
Db 2044 LNVLYETADNNTVTSYIYGSGQLLSTYENGKVF-----YHVNAGD 2087
QY 309 VQOVEFATSTSANNTA-----NP--TPAVDEIKVAKIVLSGLREGQNT 350
Db 2088 I-----IAISDSGKTVAKYQVDWANGNPKTKTEASDEVKDNRYAGYQYDEET 2135

RESULT 11
LMA_DROME
ID LMA_DROME STANDARD; PRT: 3712 AA.
AC Q00174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Laminin alpha chain precursor.
GN LANA OR LAMA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93049203; PubMed=1425586;
RA Kusche-Gullberg M., Garrison K., Mackrell A.J., Fessler L.I.,
RA Fessler J.H.;
RT "Laminin A chain: expression during Drosophila development and
RT genomic sequence."
RL EMBO J. 11:4519-4527(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RC MEDLINE=94038678; PubMed=8223265;
RA Henchcliffe C., Garcia-Alonso L., Tang J., Goodman C.S.;
RT "Genetic analysis of laminin A reveals diverse functions during
RT morphogenesis in Drosophila."
RL Development 118:325-337(1993).
RN [3]
RP SEQUENCE OF 1762-3712 FROM N.A.
RX MEDLINE=92078147; PubMed=1744083;
RA Garrison K., Mackrell A.J., Fessler J.H.;
RT "Drosophila laminin A chain sequence, interspecies comparison, and
RT domain structure of a major carboxyl portion."
RL J. Biol. Chem. 266:22899-22904(1991).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC of cells into tissues during embryonic development, and organization
CC with other extracellular matrix components.
CC -1- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING
CC STRUCTURE.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC Comprising one long and three short arms with globules at each
CC end.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT

```

```

CC MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY
CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
CC -1- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
CC DEVELOPMENT AT 10-12 HOURS.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- SIMILARITY: Contains 1 laminin N-terminal domain.
CC -1- SIMILARITY: Contains 22 laminin EGF-like domains.
CC -1- SIMILARITY: Contains 1 laminin IV domain.
CC -1- SIMILARITY: Contains 5 laminin G-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M96388; AAA28662.1; -.
DR EMBL; L07288; AAC37178.1; -.
DR EMBL; M75882; AAC28661.1; -.
DR PIR; S28399; S18253.
DR HSSP; P02468; LTLE.
DR FlyBase; FBgn0002526; LANA.
DR GO; GO:0016321; P:female meiosis chromosome segregation; IMP.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR001886; LamNT.
DR Pfam; PF00052; laminin_B_1.
DR Pfam; PF00053; laminin_EGF_18.
DR Pfam; PF00054; laminin_G_5.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00180; EGF_Lam; 19.
DR SMART; SM00281; LamB; 1.
DR SMART; SM00282; LamG; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 17.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 3712 LAMININ ALPHA CHAIN.
FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 273 332 LAMININ EGF-LIKE 1.
FT DOMAIN 333 402 LAMININ EGF-LIKE 2.
FT DOMAIN 403 447 LAMININ EGF-LIKE 3.
FT DOMAIN 448 494 LAMININ EGF-LIKE 4.
FT DOMAIN 495 540 LAMININ EGF-LIKE 5.
FT DOMAIN 541 586 LAMININ EGF-LIKE 6.
FT DOMAIN 587 631 LAMININ EGF-LIKE 7.
FT DOMAIN 632 676 LAMININ EGF-LIKE 8.
FT DOMAIN 677 731 LAMININ EGF-LIKE 9.
FT DOMAIN 732 784 LAMININ EGF-LIKE 10.
FT DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 816 1374 DOMAIN IV'.
FT DOMAIN 1375 1420 LAMININ EGF-LIKE 12.
FT DOMAIN 1421 1465 LAMININ EGF-LIKE 13.
FT DOMAIN 1466 1513 LAMININ EGF-LIKE 14.
FT DOMAIN 1514 1564 LAMININ EGF-LIKE 15.
FT DOMAIN 1565 1774 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1775 1775 LAMININ DOMAIN IV.
FT DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).
FT DOMAIN 1809 1858 LAMININ EGF-LIKE 17.
FT DOMAIN 1859 1916 LAMININ EGF-LIKE 18.

```

1917 1969 LAMININ EGF-LIKE 19.
 1970 2016 LAMININ EGF-LIKE 20.
 1971 2063 LAMININ EGF-LIKE 21.
 2084 2111 LAMININ EGF-LIKE 22.
 2111 2671 DOMAIN II AND I.
 2122 2868 LAMININ G-LIKE 1.
 2672 3048 LAMININ G-LIKE 2.
 2876 3048 LAMININ G-LIKE 3.
 3055 3233 LAMININ G-LIKE 4.
 3349 3528 LAMININ G-LIKE 5.
 3534 3709 LAMININ G-LIKE 5.
 2178 2249 COILED COIL (POTENTIAL).
 2301 2321 COILED COIL (POTENTIAL).
 2376 2450 COILED COIL (POTENTIAL).
 2541 2676 COILED COIL (POTENTIAL).
 3270 3296 POLY-THR.
 273 282 BY SIMILARITY.
 275 296 BY SIMILARITY.
 298 307 BY SIMILARITY.
 310 330 BY SIMILARITY.
 333 342 BY SIMILARITY.
 335 367 BY SIMILARITY.
 370 379 BY SIMILARITY.
 382 400 BY SIMILARITY.
 403 414 BY SIMILARITY.
 405 421 BY SIMILARITY.
 423 432 BY SIMILARITY.
 435 445 BY SIMILARITY.
 448 460 BY SIMILARITY.
 450 468 BY SIMILARITY.
 470 479 BY SIMILARITY.
 482 492 BY SIMILARITY.
 495 507 BY SIMILARITY.
 497 514 BY SIMILARITY.
 516 525 BY SIMILARITY.
 528 538 BY SIMILARITY.
 541 553 BY SIMILARITY.
 553 560 BY SIMILARITY.
 562 571 BY SIMILARITY.
 574 584 BY SIMILARITY.
 587 599 BY SIMILARITY.
 589 605 BY SIMILARITY.
 619 629 BY SIMILARITY.
 629 644 BY SIMILARITY.
 632 644 BY SIMILARITY.
 634 650 BY SIMILARITY.
 652 661 BY SIMILARITY.
 664 674 BY SIMILARITY.
 677 691 BY SIMILARITY.
 679 700 BY SIMILARITY.
 702 711 BY SIMILARITY.
 714 729 BY SIMILARITY.
 732 746 BY SIMILARITY.
 734 753 BY SIMILARITY.
 755 764 BY SIMILARITY.
 767 782 BY SIMILARITY.
 1375 1387 BY SIMILARITY.
 1377 1394 BY SIMILARITY.
 1396 1405 BY SIMILARITY.
 1408 1418 BY SIMILARITY.
 1421 1429 BY SIMILARITY.
 1423 1436 BY SIMILARITY.
 1438 1447 BY SIMILARITY.
 1450 1463 BY SIMILARITY.
 1466 1480 BY SIMILARITY.
 1468 1487 BY SIMILARITY.
 1489 1498 BY SIMILARITY.
 1501 1511 BY SIMILARITY.
 1514 1526 BY SIMILARITY.
 1516 1533 BY SIMILARITY.
 1535 1544 BY SIMILARITY.
 1547 1562 BY SIMILARITY.
 1559 1574 BY SIMILARITY.
 1581 1595 BY SIMILARITY.
 1611 1624 BY SIMILARITY.
 1626 1640 BY SIMILARITY.
 1642 1656 BY SIMILARITY.
 1658 1672 BY SIMILARITY.
 1674 1688 BY SIMILARITY.
 1690 1704 BY SIMILARITY.
 1706 1720 BY SIMILARITY.
 1722 1736 BY SIMILARITY.
 1738 1752 BY SIMILARITY.
 1754 1768 BY SIMILARITY.
 1770 1784 BY SIMILARITY.
 1786 1800 BY SIMILARITY.
 1802 1816 BY SIMILARITY.
 1818 1832 BY SIMILARITY.
 1834 1848 BY SIMILARITY.
 1850 1864 BY SIMILARITY.
 1866 1880 BY SIMILARITY.
 1882 1896 BY SIMILARITY.
 1898 1912 BY SIMILARITY.
 1920 1934 BY SIMILARITY.
 1936 1950 BY SIMILARITY.
 1952 1966 BY SIMILARITY.
 1968 1982 BY SIMILARITY.
 1984 1998 BY SIMILARITY.
 2000 2014 BY SIMILARITY.
 2016 2030 BY SIMILARITY.
 2032 2046 BY SIMILARITY.
 2048 2062 BY SIMILARITY.
 2064 2078 BY SIMILARITY.
 2080 2094 BY SIMILARITY.
 2096 2110 BY SIMILARITY.
 2112 2126 BY SIMILARITY.
 2124 2138 BY SIMILARITY.
 2140 2154 BY SIMILARITY.
 2156 2170 BY SIMILARITY.
 2172 2186 BY SIMILARITY.
 2188 2202 BY SIMILARITY.
 2204 2218 BY SIMILARITY.
 2220 2234 BY SIMILARITY.
 2236 2250 BY SIMILARITY.
 2252 2266 BY SIMILARITY.
 2268 2282 BY SIMILARITY.
 2284 2298 BY SIMILARITY.
 2300 2314 BY SIMILARITY.
 2316 2330 BY SIMILARITY.
 2332 2346 BY SIMILARITY.
 2348 2362 BY SIMILARITY.
 2364 2378 BY SIMILARITY.
 2380 2394 BY SIMILARITY.
 2396 2410 BY SIMILARITY.
 2412 2426 BY SIMILARITY.
 2428 2442 BY SIMILARITY.
 2444 2458 BY SIMILARITY.
 2460 2474 BY SIMILARITY.
 2476 2490 BY SIMILARITY.
 2492 2506 BY SIMILARITY.
 2508 2522 BY SIMILARITY.
 2524 2538 BY SIMILARITY.
 2540 2554 BY SIMILARITY.
 2556 2570 BY SIMILARITY.
 2572 2586 BY SIMILARITY.
 2588 2602 BY SIMILARITY.
 2604 2618 BY SIMILARITY.
 2616 2630 BY SIMILARITY.
 2632 2646 BY SIMILARITY.
 2648 2662 BY SIMILARITY.
 2664 2678 BY SIMILARITY.
 2680 2694 BY SIMILARITY.
 2696 2710 BY SIMILARITY.
 2712 2726 BY SIMILARITY.
 2728 2742 BY SIMILARITY.
 2744 2758 BY SIMILARITY.
 2760 2774 BY SIMILARITY.
 2776 2790 BY SIMILARITY.
 2792 2806 BY SIMILARITY.
 2808 2822 BY SIMILARITY.
 2824 2838 BY SIMILARITY.
 2840 2854 BY SIMILARITY.
 2856 2870 BY SIMILARITY.
 2872 2886 BY SIMILARITY.
 2888 2902 BY SIMILARITY.
 2904 2918 BY SIMILARITY.
 2920 2934 BY SIMILARITY.
 2936 2950 BY SIMILARITY.
 2952 2966 BY SIMILARITY.
 2968 2982 BY SIMILARITY.
 2984 2998 BY SIMILARITY.
 3000 3014 BY SIMILARITY.
 3016 3030 BY SIMILARITY.
 3032 3046 BY SIMILARITY.
 3048 3062 BY SIMILARITY.
 3064 3078 BY SIMILARITY.
 3080 3094 BY SIMILARITY.
 3096 3110 BY SIMILARITY.
 3112 3126 BY SIMILARITY.
 3124 3138 BY SIMILARITY.
 3140 3154 BY SIMILARITY.
 3156 3170 BY SIMILARITY.
 3172 3186 BY SIMILARITY.
 3188 3202 BY SIMILARITY.
 3204 3218 BY SIMILARITY.
 3216 3230 BY SIMILARITY.
 3232 3246 BY SIMILARITY.
 3248 3262 BY SIMILARITY.
 3264 3278 BY SIMILARITY.
 3280 3294 BY SIMILARITY.
 3296 3310 BY SIMILARITY.
 3312 3326 BY SIMILARITY.
 3324 3338 BY SIMILARITY.
 3340 3354 BY SIMILARITY.
 3356 3370 BY SIMILARITY.
 3372 3386 BY SIMILARITY.
 3388 3402 BY SIMILARITY.
 3404 3418 BY SIMILARITY.
 3420 3434 BY SIMILARITY.
 3436 3450 BY SIMILARITY.
 3452 3466 BY SIMILARITY.
 3464 3478 BY SIMILARITY.
 3480 3494 BY SIMILARITY.
 3496 3510 BY SIMILARITY.
 3512 3526 BY SIMILARITY.
 3524 3538 BY SIMILARITY.
 3540 3554 BY SIMILARITY.
 3556 3570 BY SIMILARITY.
 3572 3586 BY SIMILARITY.
 3588 3602 BY SIMILARITY.
 3604 3618 BY SIMILARITY.
 3620 3634 BY SIMILARITY.
 3632 3646 BY SIMILARITY.
 3648 3662 BY SIMILARITY.
 36

DR SGD: S0001431: YIL169C.
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR PROSITE: PS00111; CHROMOTAXIS_TRANSLOC_2; 1.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 995
 FT DOMAIN 26 253
 FT CARBOHYD 28 28
 FT CARBOHYD 35 35
 FT CARBOHYD 468 468
 FT CARBOHYD 664 664
 SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match
 Best Local Similarity 6.4%; Score 128; DB 1; Length 995;
 Matches 73; Conservative 100; Mismatches 152; Indels 128; Gaps 15;

QY 4 SITKRDANPNNGOTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTNN 63
 DB 136 SVSQSSASDVSSVSQSSSSASDVSSVSQSSASD-----VSSVSQSSASSTD 188
 QY 64 LNATLEQLKMAKTNIESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLST 123
 DB 189 VSSVSQSSASDVSSVSQSSSSASDVSSVSQSSASSTDVSSVSQSSASSTGVSSS 248
 QY 124 AYNQIRNLVLDYNKASSLITKTLPLNGTLLDSNEITTTANKNNITLSTINEQKTNAD 183
 DB 249 GSQSVSS-----ASGSSSSFPQST-----SSASTASGSATSNLSLSTSSASAS 293
 QY 184 ALSNSFIKK-----VIONNEQSF-----VG 203
 DB 294 AFASNSLSSDGTIYLPPTTISGDLTLGKVTATGEVVVAGAKITLDDGKYSFADLK 353
 QY 204 TETNANVOPSNYSFVAFSADVPVNYKYARRTVWNGDE---PSSRLANTNSITDVSMTY 260
 DB 354 VYGDLLVKSKETYPGTEEDISEGFDVTGN--FNAESASATSASIYFSPFSFNSGDI 411
 QY 261 SLA-GTNTKYQSFNYPSTGYLYPPYKLVKAADANNVGLQYKLNNGNVQVQVEPATST- 318
 DB 412 SLSLSKSKKGEVTFSPYNSGAFSP-----SNAL-----LNGGSVSGLQRRDDTE 456
 QY 319 -SANNNTANPTPAVDEIKVAKIVLSLREGQNTIELSVPTGKNNKVPAMIGNIYL--- 374
 DB 457 GSVNGEIN-----LNGSTYIVPEVSGKGTVNIIS---GNLYLHP 496
 QY 375 -----SSNENNADKIP--GY 387
 DB 497 DFTGTQTVYKGEGLVAVDPTETNATPIPVVGY 529

RESULT 13
 FLGK_BUCAP
 ID FLGK_BUCAP STANDARD; PRT; 544 AA.
 AC Q8K9K0.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flagellar hook-associated protein 1 (HAP1).
 GN FLGK OR BUG334.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22084549; PubMed=12089438;
 RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
 RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
 RT *50 million years of genomic stasis in endosymbiotic bacteria.*;
 RL Science 296:2376-2379(2002).
 CC -1- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
 CC FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: AE014109; AAM67888.1; -
 DR InterPro: IPR001444; Flag_bb_tod.
 DR PRINTS: IPR002371; Flag_hoopa1.
 DR PROSITE: PR01005; FLGHOOKAP1.
 DR PROSITE: PS00588; FLAGELLA_BB_ROD; FALSE_NEG.
 KW Flagella; Complete proteome.
 SQ SEQUENCE 544 AA; 63317 MW; 439E8AD7781363CB CRC64;

Query Match
 Best Local Similarity 6.3%; Score 127; DB 1; Length 544;
 Matches 97; Conservative 54; Mismatches 165; Indels 106; Gaps 21;

QY 16 QTQLEAARMELTDLINAKA--MTLASLDYAKIEASLSAYSEA--ETVNNNLNATLEQL 71
 DB 82 QTKIE-EVLKLEDLFGKSNIFTLLINOLYSSIENDIVNNHGVFNENIENLNTIIHEL 140
 QY 72 KM-----AKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 120
 DB 141 KNFDEKLSFLERDVKESIKERIKQAN-----ILINKIYDINIDIR-----F 181
 QY 121 SSTAYNQIRNLVLDYNKASSLITKTLPLNGTLLDSNEITTTANKNNITLSTINE-QK 179
 DB 182 PPTA--QIPNRIDSIDKRDKLIDE-LNDIIQVKIKEN--STFKVCLNNGCIIDDDYNK 236
 QY 180 TNADALSNSFIKKIQNNEQSFVGTFTNANVOPSNYSFVAFSADVPVNYKYARRTVWNG 239
 DB 237 KNLMTLSTDDDKYIS-----VGYFDDNEQRLKKEHWPASISGAL-LTFRREDLQNA 289
 QY 240 DEPSRRIAN-TNSTDYSWI-YSLAGNTKYQFSPFN-----YGPSTGYLYPPYKLVKA 292
 DB 290 KNIKIGLTINFADSNVHTLGYDILGNIGKQVFRKISNPEIISSSQNSQSLPTSIRKWD 349
 QY 293 ADA-----NNVGLQYKLNNGNVQVQVEPATSTANNTANPTPAVDEIKVAKIVLS 342
 DB 350 ADAQNTNIVFLKNNHWTVRLRHSV--VEPDIIQQDNNNT-----YITFD 393
 QY 343 GLRFQNTIELSVPTGEGNNKVP-----MIGNIYLS-----SNENNAD 382
 DB 394 GIEF-----KIBGNDAGNNYMKPYSTLTNLELLIINKDLFSISSSSDDLQKNRNNAI 448
 QY 383 KI 384
 DB 449 KI 450

RESULT 14
 YEF3_YEAST
 ID YEF3_YEAST STANDARD; PRT; 956 AA.
 AC P32618;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 106.1 kDa protein in GLY1-GDA1 intergenic region.
 GN YEL043W OR SYGP-ORF14.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,

RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL Nature 387:78-81(1997).
RN [2]
RP FIBRONECTIN TYPE III DOMAIN.
RX MEDLINE=97148176; PubMed=8994808;
RA Bateman A., Chotha C.;
RT "Fibronectin type III domains in yeast detected by a hidden Markov
model.";
RL Curr. Biol. 6:1544-1546(1996).
CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18779; AAB64999.1; -;
DR PIR: S30834; S30834.
DR SGD: S0000769; YEL043W.
DR InterPro: IPR003961; FN_III.
DR SMART: SM00060; FN3; 1.
KW Hypothetical protein.
FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
FT SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;
SQ
Query Match 6.3%; Score 127; DB 1; Length 956;
Best Local Similarity 19.7%; Pred. No. 6.3;
Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;
QY 5 ITKDPNNGQTOLEARMELDTL-----INA-KAMTLASLDQVAKIEASLSAYSEA 57
DB 337 LNSASVANINKEIESIQNEISKEESNRKLNASKSLTISVIVNANVENDPIASGEL 396
QY 58 ETVNNIN-ATLE-----QLKMAKTNLESAINOANTDKTTFDNEHNLVYAYKALK 107
DB 397 SAVLKLENDFTLEKNGFLSNAGEFLSKLNADSLIKMIKQELSIOQE---LEANWKLQR 453
QY 108 TLEQRTATNLEGLSTAYNQIRNLVDLYNKASLIKTL-DPL--NGFTLLDSNETTA 164
DB 454 SNLLKRTSALE-----NQF--NEMSLNR--NLKTKLMVQPKYKNGDSLAATNSNSA 502
QY 165 NKN-----INNTLSTNEQKTNADAL-----185
DB 503 EKNRSSGSIQLPLSNNSRTGSDILSNKNSIN--NSNADSAAPPLRLHNPVSPSPNEP 560
QY 186 ---SNSPIKKVQIYQNE--QSPVGTFTNAN---VQPSNYSFVA-----FSADVTPVNYK 230
DB 561 IQPSSLSQLTQTDNRSLNLSHNSNENKQOPSSYSHALPTTANATATATATN-G 619
QY 231 YARTVTVNG---DEPSSRILA-----NTNSITDVSWIYSLAGTNTKYQPSFSNYG 277
DB 620 HSRSLNLTWTAQAFQPSHQQVSTELDQAFEDVNDANHL--ISGLQNMIVDEYDIPDNISNY 677
QY 278 -----PSTGYLFFPKYKLVKAADANVGVQYKLNNGNVQ 310
DB 678 KGFTTDELNDVYWKQPOVRSTNESLFTGTGTPASSYK-----ANPVISPY--SSSHLR 729
QY 311 QVERATSTANNTANTPTPAVDKTKVAKIVLSGLRF-----GQNTIELS-- 354
DB 730 QTSNATNTNPMHQPSLLAATLNDPSLQSFVRSRGSFYSAOPANSLQNNINGNETNISPR 789
QY 355 -----VPTGEGNMKVAPMIGNIYVLSNENNAADKI 384
DB 790 ISSDFNLLVPLNPLSLNDVPIVPCNNTTLTPSHSNILTMNHQPTADNI 838

RESULT 15
PST1_YEAST
ID PST1_YEAST STANDARD; PRT; 444 AA.
AC Q12355;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protoplast secreted protein 1 precursor.
GN PST1 OR YDR055W OR D4214 OR YD9609.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96381250; PubMed=8789263;
RA Brandt P., Ramlow S., Otto B., Bloecker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
of Saccharomyces cerevisiae chromosome IV.";
RL Yeast 12:85-90(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION.
RX MEDLINE=99251092; PubMed=10234784;
RA Pardo M., Montecilla L., Pla J., Sanchez M., Gil C., Nombela C.;
RT "Two-dimensional analysis of proteins secreted by Saccharomycetes
cerevisiae regenerating protoplasts: a novel approach to study the
cell wall.";
RL Yeast 15:459-472(1999).
RN [4]
RP GPI-ANCHOR.
RX MEDLINE=20469049; PubMed=11016834;
RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
RT "Up-regulation of genes encoding glycosylphosphatidylinositol
(GPI)-attached proteins in response to cell wall damage caused by
disruption of FKS1 in Saccharomycetes cerevisiae.";
RL Mol. Gen. Genet. 264:64-74(2000).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
SECRETED BY REGENERATING PROTOPLASTS.
CC -1- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X84162; CAA58971.1; -;
DR EMBL: Z74351; CAA98873.1; -;
DR EMBL: Z49209; CAA89084.1; -;
DR PIR: S54039; S54039.
DR COMPUYEAST-2DPAGE; Q12355; -;
DR SGD: S00024627; PST1.
DR GO: GO:0009277; C:cell wall (sensu Fungi); IDA.
DR InterPro: IPR000494; EGFR_L_domain.
DR Pfam: PF01030; Recep_L_domain; 1.
KW Glycoprotein; Membrane; GPI-anchor; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 ? PROTOPLAST SECRETED PROTEIN 1.
FT PROPEP ? 444 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 356 416 SER-RICH.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).

Search completed: October 8, 2003, 17:30:01
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:26:01 ; Search time 42 Seconds
(without alignments)
902.154 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086
Perfect score: 2004
Sequence: 1 GCMSITKDKANPNNGQIQLE.....SSNENNADKIPCYRRPGTFL 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	41.5	702	2 S48754	major surface prot
2	754	37.6	702	2 S48753	major surface prot
3	744	37.1	650	2 S48751	major surface prot
4	743	37.1	649	2 S48752	major surface prot
5	713	35.6	647	2 A49218	hemagglutinin homo
6	575.5	28.7	386	2 S48755	major surface prot
7	189.5	9.5	320	2 S31560	major surface prot
8	175	8.7	6713	2 B99921	hypothetical prote
9	173	8.6	1302	1 JG6009	surface-located me
10	165.5	8.3	661	2 AG2422	FmtB protein [impo
11	164.5	8.2	2481	2 D90011	hypothetical prote
12	163.5	8.2	3890	2 C89921	hypothetical prote
13	159	7.9	933	2 S41539	fibrinogen-binding
14	156	7.8	807	2 B71605	hypothetical prote
15	155	7.7	1072	2 A86827	hypothetical prote
16	155	7.7	4688	2 F82885	hypothetical prote
17	154.5	7.7	682	2 S44131	subtilisin-like pr
18	154.5	7.7	1365	2 T30822	lmpI protein - Myc
19	151.5	7.6	1237	2 D71850	probable outer mem
20	151	7.5	624	2 PG6003	surface membrane p
21	150.5	7.5	1051	2 T18351	lmpI protein - Myc
22	145	7.2	2399	2 H71879	toxin-like outer m
23	144.5	7.2	1487	2 AG2560	hypothetical prote
24	144	7.2	1107	2 AC0976	probable autotrans
25	143.5	7.2	751	2 T40462	ser-lys rich hypot
26	142	7.1	135	2 B49218	hemagglutinin homo
27	142	7.1	2660	2 E85822	probable invasiv z
28	141.5	7.1	1645	2 F96907	phage-related prot
29	140.5	7.0	3194	2 D71917	toxin-like outer m

30	139.5	7.0	2401	2 T28676	hoptry protein -
31	138.5	6.9	1314	1 TBYR6	transcription regu
32	137.5	6.9	820	2 T17519	cell surface antig
33	137.5	6.9	926	2 AE1130	conserved hypotet
34	137.5	6.9	989	2 D89852	fibrinogen-binding
35	137.5	6.9	1524	2 S68553	surface layer prot
36	137.5	6.9	4152	2 T31102	filamentous hemagg
37	137	6.8	719	2 S55119	hypothetical prote
38	137	6.8	5005	2 F82884	hypothetical prote
39	136.5	6.8	568	2 E97066	membrane associate
40	136.5	6.8	1073	2 S14032	kinesin-related pr
41	136.5	6.8	2271	2 F90073	hypothetical prote
42	135.5	6.8	1085	2 T38378	kinesin-like prote
43	135.5	6.8	1238	2 A64596	hypothetical prote
44	135	6.7	2269	2 T28677	hoptry protein -
45	134.5	6.7	589	2 B97806	hypothetical prote

ALIGNMENTS

RESULT 1

S48754
major surface protein (clone pmGal.4) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48754
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48754
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-702 <WAR>
A:Cross-references: EMBL:L28424
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 41.5%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 3.3e-38;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY	9	DANPNNG-----OTOLEAARMELDLINAKAMTILASLODYAKIBASISAYSEATVNN	62
DB	68	NTNPGNGGGTDNAAQQAALAAKKELSDLLATQNSLSTYADYANIQTTLTAAYTTAKRSTSD	127
QY	63	NLNATLEOLKMAKTNLESAINQANTDKTTFDNEHNPVLEAYKALKTTLEQRAINLEGLSS	122
DB	128	NTSATLEQVKTSATSLQTAIDTAASRSTSFDEKNPELIKAYALKETLAKWRNSLSGLTD	187
QY	123	TAYNQRINRLVDLYNKASLIKTLDPLNGGTLSDSNETTANKNNNTLSINEQKTNWA	182
DB	198	SNAFKIKNTALYOSGKDVIYTKLDPLM-GTAINLSAVSQANTHISNAVSKLEWTKTNWA	246
QY	183	DALSNSFKIKVQNNEQSFVGFTTANVQPSISVFASADVTP-----VNYKYARRTV	236
DB	247	TVLATSFVKELVKNKLTGIDT-TNNOEQPGNYSFVGYSVDVTTGSDNARPKNWSAQKRV	305
QY	237	WNGD-----EPSSRLIANTNSITDVSWISYSLAGTNTKYQSFNSGPGSYFYPPKLVK	291
DB	306	WTSNTDILSQPQAEENGQOASDPYSWIYNLGMAKAYSLTFTNYIGPSTGFLYFFYKLVN	365
QY	292	RADANNVGLQYKLNNGNVQOEFATS-----TSANN-----TTANTPTPAVD	332
DB	366	SSDSKVALEYKLNESAVKTIIDFSQSPVAVSDATRENNRSTATAPOAGSTSTEINAPTLD	425
QY	333	EIKVAKIVLSGLRFQONTIELSVPT-TEGNNKNKVPAMIGNIYLSNENNAKD	383
DB	426	DIKIAKIVLSNLKFGSNTIEFSVPTAKEGTSKVAPMIGNMYLTSSDRDVK	477

RESULT 2

S48753
major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48753
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface p
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-702 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.6%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 5.7e-34;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;
QY 11 NPNGQT----QLEAARMETDLINAKAMTASLODYAKIEASLSAYSEAEVNNLNA 66
DB 69 NPNGNTTPBQQAARAKTLTDLGTENTVALYADYAKIQSTLSAYMTAKTASENTSA 128
QY 67 TLEQKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEQRTNLEGLSTAYN 126
DB 129 TLENRSASTLQAADKAANDKRVDSVNOPLVAAYNNLKTTLKSTTSLEGLSENKYG 188
QY 127 QIRNLVDLYNKASSLITKTLPLNG--GTLDSNEITANKNINNTLS--TNEQKTN 182
DB 189 GIKNHLKSLFDTGSAITAKTLDTSGRPTLEKVN-----ANNGIKMAISPESLKKWKGNA 245
QY 183 DALSNSFIKKVIONNEQSFVGTNTANVQPSNYSEVAFSADVTP-----VNYKVARVTW 237
DB 246 DKF-NEPEKNPLSKKLUKSTDAHQEQPANFSAISYVDLTSNQLNPNWFAQRKVV 304
QY 238 NGD--EPSSRLANTN--SIITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAAD 294
DB 305 TSENQPGKTLAVSSPSVATDVSWIYSLAGEKTYLTFEYGYGPDNAFLYPLKLVKAAD 364
QY 295 ANVGLQYKLVNGNVQVEF-----ATST-----SAN 321
DB 365 SSSVALQYSLNKTSSKLNFKPAETVSTNQDSNEVATTSTTEARSYKVLVADEAATS 424
QY 322 NTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAPMIGNIYSSNENNA 381
DB 425 NNMNHTPTVSDINIAKVTLSGLTFGENTIEFSVPEG----KVAPMIGNMYLTSNSESQ 479
QY 382 DK1 384
DB 480 VK1 482

RESULT 3

S48751
major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum
N:Alternate names: major hemagglutinin pmGA
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48751; A44793
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface p
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48751
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-650 <MAR>
A:Cross-references: GB:U90714; EMBL:L28423; NID:g1905869; PIDN:AA50152.1; PID:g1905870
R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whithear, K.G.

Infect. Immun. 60, 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepti
A:Reference number: A44793; MUID:92363591; PMID:1379991
A:Contents: S6
A:Accession: A44793
A:Status: preliminary
A:Molecule type: protein
A:Residues: 26-42 <MAR>
A:Note: sequence extracted from NCBI backbone (NCBIP:111017)
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.1%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 1.8e-33;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;
QY 11 NPNGQ-----TOLEAARMETDLINAKAMTASLODYAKIEASLSAYSEAEVNNN 63
DB 51 NPGDGGGMNAAQAELAAARMGLTTFVDSKAKNLGYDYKKTONTLTAKYDAKTVLDN 110
QY 64 LNATLEQKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEQRTNLEGLSST 123
DB 111 SSSITQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYADA 170
QY 124 AYNOIRNLVDLYNKASSLITKTLPLNGTGLDLSNEITANKNINNTL--STINEQKTN 181
DB 171 QYAGIKMHLISGLYDAGAKAITTKLEPEVGP-LTAGAVTMAINTKIYEAIKDEVLPNPKKN 229
QY 182 ADALSNSFIKKVIONNEQSFVGTNTANVQPSNYSEVAFSADVTPV-----NKKYAR 233
DB 230 AYKLADSFYKQVLVKEKTIQVEAHN-KAQPANYSFVGYSDITGTANGQTSIPNNYQAQ 288
QY 234 RTVW-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLY 284
DB 289 RTITFGDEP--RSVSNTPVDGOTMAQPLSNVSWIYSLAGTGAKYTLEFTYYPSTGYLY 346
QY 285 FPKLVKKAADANNVGLYKLVNGNVQVEFATSTSANNTAN-PTPAVDEIKVAKIVLSG 343
DB 347 FPKLVNTSDQMKLGLEYKLNDL-----TEPSAITFGNEQTMNGKTPVDINAVKVTLAN 402
QY 344 LRFQNTIELSVPTGEGNNKVPAPMIGNIYSSNENNAKDI 384
DB 403 LIFGSKNIEPSVPA-----EKVSPMIGNMYLSSSPNNWNKI 438

RESULT 4

S48752
major surface protein (clone pmGAL.2) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48752
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surfac
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48752
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-649 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 2e-33;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;
QY 9 DANPNNGQ-----TOLEAARMETDLINAKAMTASLODYAKIEASLSAYSEAEVTVN 61
DB 48 DTNPGDGGGMNAAQSLEAARMGLTTFIDSKAKNLGLYDYKKTONTLTAKYDAKTVL 107

Qy 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAKKALKTTLLEQRATNLEGLS 121
Db 108 DNSSSTQNLNEAKTRLETAIRTAATSKOTFDEQHAELVYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNRLVDLYNKASSLITKTLPLNGGTTLLDSNEITANKNNNTL--STINEQK 179
Db 168 DAQYAGIKMHLGLYDAGKAITTLEPEVGGDP-LTASAVMMATKIVEALKDEVLPQK 226
Qy 180 TNADALNSFIKKYQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVP-----NYKY 231
Db 227 ENATKLADSFVKQVLVKEKITGVBEAHN-RAQPANYSFVGYSVDITGTTGOTSIPNWDY 285
Qy 232 ARRTVW-NGDEPSSRIANT-----NSTIDVSWIYSLAGTNTKKQFSFNYPSTGY 282
Db 286 AQRITFTNSDEP--RSISNTPADQGTMAQPLSNVSWIYSLAGTCAKTTLETTYGPTGY 343
Qy 283 LYFFPKLVKAADANNVGLQKLNNGVQVFEATSTSA-----NNTTANPTTPAYDEIKVA 337
Db 344 LYFFPKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPPTVNDINVA 395
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 384
Db 396 KVTILANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437

RESULT 5
A49218
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin
A:Reference number: A49218; MUID:93162830; PMID:8432610
A:Accession: A49218
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-647 <MAR>
A:Cross-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIPI:125183)
C:Genetics:
A:Genetic code: SGC3

Query Match 35.6%; Score 713; DB 2; Length 647;
Best Local Similarity 41.1%; Pred. No. 8.9e-32;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

Qy 9 DANPNNG-----QTLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
Db 48 DTNFGDGGGMMAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTTRAYDAAKTVL 107
Qy 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAKKALKTTLLEQRATNLEGLS 121
Db 108 DNSSSTQNLNEAKTRLETAIRTAATSKOTFDEQHAELVYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNRLVDLYNKASSLITKTLPLNGGTTLLDSNEITANKNNNTL--STINEQK 179
Db 168 DAQYAGIKMHLGLYDAGKAITTLEPEVGGDP-LTASAVMMATKIVEALKDEVLPQK 226
Qy 180 TN-----ADALNSFIKKYQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVP----- 227
Db 227 ENATKLADSLSSIVKKITGVEE-----AHNKAQAPANTSFVGKRWYTELLELDKQVFP 279
Qy 228 NYKARTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNTKKQFSFNYPSTGY 278
Db 280 NWDYAQRTIFTNSDEP--RSISNTPADQGTMAQPLSNVSWIYSLAGTCAKTTLETTYGPTGY 337
Qy 279 STGYLYFPYKLVKAADANNVGLQKLNNGVQVFEATSTSA-----NNTTANPTTPAYDE 333
Db 338 STGYLYFPYKLVNTSDQVKLEYKLN-----ATKPSAITFGSDQTMNGKTPPTVND 389

Qy 334 IKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 384
Db 390 INVAKVTILANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 435
RESULT 6
S48755
major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48755
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48755
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-386 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692
A:Note: The sequence of residues 385-386 and the corresponding nucleotide sequence ar
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 28.7%; Score 575.5; DB 2; Length 386;
Best Local Similarity 42.3%; Pred. No. 1.5e-24;
Matches 138; Conservative 50; Mismatches 113; Indels 25; Gaps 10;
Qy 9 DANPNNG-----QTLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 62
Db 64 NTNFGGGGMDNSAQALAAARKEUILLATONSLSLTADYAKIKNDITAAATTAETASQ 123
Qy 63 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAKKALKTTLLEQRATNLEGLSS 122
Db 124 NQAATLEQVKNAASTLQTAINTAVNEKKVFDENNSSELVATNLKTTLEGENTLLAAAND 183
Qy 123 TA-YNQIRNRLVDLYNKASSLITKTLPLN-GGTTLLDSNEITANKNNNTL--STINEQ 178
Db 184 SANTYGIKTHLLSYNQAKTITSTL--LNDAGOSPKNDMVVKINKEITDAINPTLLNQ 241
Qy 179 KTNADALNSFIKKYQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVP-----VNY 229
Db 242 KANADMLATSTFKQVL-NDAOLTSGSSETSQTOPQPPGNISFVGVSDVDTTGSNNARNPW 300
Qy 230 KYARTVWNGDEPSSRIANTNSITDVSWIYSLAGTNTKKQFSFNYPSTGYLYFPYK 288
Db 301 NFAQRKYWDTRAPLAQTEQSNKLTDSWIYSLSCMGAKYVTFDYGCASNAYLYFPYK 360
Qy 289 LVKAADANNVGLQKLNNGVQVFEF 314
Db 361 LVQTNDD--NVGLQVYLVNNTTPKLVNF 384

RESULT 7
S51560
major surface protein (clone pmGAL.6) - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S51560; S48757
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S51560
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-320 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62415.1; PID:g535688
C:Genetics:
A:Genetic code: SGC3

Query Match 9.5%; Score 189.5; DB 2; Length 320;

[illegible]

Db 3836 AQNAVEAALQRVNTAKDALNGDAKLI 3861

RESULT 13

S41539

N;Alternative names: Staphylococcus aureus

C;Species: Staphylococcus aureus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999

C;Accession: S41539; S36630

R;McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus

A;Reference number: S41539; MUID:94224142; PMID:8170386

A;Accession: S41539

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-933 <MCD>

A;Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match 7.9%; Score 159; DB 2; Length 933;

Best Local Similarity 22.8%; Pred. No. 0.3;

Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY 3 MSTTKKDANPNQQTOL--EARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEIV 60

DB 73 VSDTKTSSNTNGETSAQNPAAQQTQSSTNATT-----EETPTVGEATTT 120

QY 61 NNINLATLEOLKMAKTNLESAINQAQNTDPTFNEHPNLVEAYKALKTTLEQKATNLEGL 120

DB 121 TTQANTPATQSSNTNAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172

QY 121 SST-----AYNQIRNNLVLYNK-----ASSLITKTLDPPLNGTLL 156

DB 173 STTQDTSTEAPTSNNESAPQSTDAKNVQVAVNTSAPRMRAFSLAADAAAPAAAGTDI 232

QY 157 DSNEITTANKNINNTLSTNEQ---KTN-ADALSNSFIK---KVLQNEQSFVGTFTN 207

DB 233 -TNQNTVTVGIDSGTIVPHQAGYKLNLYGFSVPNSAVKGDFTFKITVPKELNLTGVT 291

QY 208 ANVQPSNYSEAFVADPTVNYKYARTVWNGDEPSSRIILANTNSITDVSIIYSLAG-TN 266

DB 292 AKYPP-----IMAGDQ-----VLANGVIDSDGNVITFTDYVN 324

QY 267 TKYQFSNPGSPGYLYFYKLVKAADANNVGLQKLNNGVQVQVEAFATSTANNNTAN 326

DB 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTAN 362

QY 327 PTPAVDEIKVAKI-----VLSGLRPGQNTIELSVPTGEGNMKVAPMI-GNIIYSSNE 378

DB 363 KTVLVDIEYKGFNLSIKGTIDQIDKNTNYRTQIIVNPSGDNVIAPLVTGNLKPWDS 422

QY 379 N 379

DB 423 N 423

RESULT 14

B71605

hypothetical protein PF08050c - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

C;Accession: B71605

R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.

Science 282, 1126-1132, 1998

A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A;Reference number: A71600; MUID:99021743; PMID:9804551

A;Accession: B71605

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-807 <GAR>

A;Cross-references: GB:AE001420; GB:AE001362; NID:g3845287; PIDN:AAC71959.1; PID:g384529

A;Experimental source: clone 3D7

C;Genetics:

A;Gene: PF08050c

Query Match 7.8%; Score 156; DB 2; Length 807;

Best Local Similarity 23.8%; Pred. No. 0.36;

Matches 73; Conservative 48; Mismatches 124; Indels 62; Gaps 14;

QY 60 VNNLNATLEOLKMAKT-----NLESAINQA--NTDK-----TTFDNEHPNLVEAYKALKT 108

DB 468 INNISYT--QLKMNINFINHESPINQHNNTFKVNNNDINFTFETPKMKKKKEKK 525

QY 109 TLEQKATNLEGLSSAYAYNQIRNNLVLYNKASSLITKTLDPGLNGTLLDSNEITANKNI 168

DB 526 NIHFNNNNNNNNKCLYKIDINQN---DHNNSTINTNQNFEDI-----NNVKTEONL 574

QY 169 ---NNTLSTINEQKTNADALSNSFIKKVIQNEQSFVGTFTNANVQPSNYSEAFSADVT 225

DB 575 QKKHKMSQVSKQSNKNNKNSHLKKQININ-----TNNMDKNNSHISKNVIVD 626

QY 226 PVNYK-----YARRTVWNGDEPSS-----RILANTNSITDVSIIYSLAGTNTKYQ-FSPSN 275

DB 627 DNKLKSSHADNSNEIVTKGKKKNTKKKKINNINSVNNVNNINSNNINSMNNIISMNN 686

QY 276 YGPSTGYLYFPYKLVKAADANNVGLQYK-----LNNGNVQVQVEFATSTSAN 321

DB 687 VNNMNPMTFFPNVNIQ--KDDSNIALLYNNKPIDFNFNQNLNHNHMIQNNTNNTNVMNLN 745

QY 322 N--TTAN 326

DB 746 NNLTTSN 752

RESULT 15

A86827

hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL140

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C;Accession: A86827

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: A86827

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1072 <STO>

A;Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: yqfG

Query Match 7.7%; Score 155; DB 2; Length 1072;

Best Local Similarity 21.0%; Pred. No. 0.59;

Matches 97; Conservative 76; Mismatches 154; Indels 136; Gaps 20;

QY 9 DANPNNGOTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSE-AETVNNNL--- 64

DB 430 ESNVSVDQTSSEAS--TNSNSSISLSPNISSTSD--SESATNSDSFSNVAEVANNSLASV 486

QY 65 -NATLEOLKMAKTNLESAINQAQNTDPTFNEHPNLVEAYKALKTTLEQKATN----- 116

DB 487 NNSSSSVLSTSTADNLGINQSGSDNLTKDSSEISTSGAFLSSNQTSSEASTNSSSISL 546

QY 117 -----LEG-LSTAYNQI-----RNNLVLYNKASSLI--TKTLDPL-----NGG 153

DB 547 SPNSISSTVLESTTSSTNSFNVAEVANNSLVSSSSSVLSTSTADNLEINQFGSDN 606

QY 154 TLDSDNEITTANKNINNTLSTINEQKTNADALSNSFIKKVIQNEQSFVGTFTNANVQPS 213

DB 607 LTQDSSEISTSG-----AFLSNQTSSEASSNS-----MSSINSPSLSLSTFNSSEAT 655

QY 214 NYSEAFSADVTVPYNYKYARTVWNGDEPSSRI-----ANTNSITDVSIIYSLAG 264

[illegible]

Search completed: October 8, 2003, 17:32:35
Job time : 44 secs

THIS PAGE BLANK (USPTO)

Query Match	6.28;	Score 124.5;	DB 6;	Length 1010;	
Best Local Similarity	21.28;	Pred. No. 7.8;			
Matches	98;	Conservative	87;	Mismatches 147; Indels 131; Gaps 25;	
Qy	3	MSITKKD----	ANPNNGOTOLEAA--RMEITDLINAKAM--	TLASLDQVAKTEASISLSAYS 55	
Db	402	MDITEQPCP	AGNOODGLOKREDIMDISVIRSPAVNDTMVQSPARVVKICANSII	461	
Qy	56	EAE-----		TVNNLNATLEOLKMAKTNLESAINOAN 86	
Db	462	DSOKSIVFG	DEMSIDETQNDGTLPLPKSNVEVTTNDVYTSLREQEENASENVSWINESS	521	
Qy	87	T----	DKTTPDNEHPNLVEAKA--LKTTLEQRATNLE--	GLSSFAYNQIIRNLNLDLYNKA 139	
Db	522	VHSIDKKSF-----	MLTEERAEPMHSDMDVAOKLEDGDSKTP-----	VILASQS 568	
Qy	140	SSLITK-----	TLD--PLNGOTLSDNSNITTANKNINNTLSTINEQ----	KYNAD 183	
Db	569	ASLTKEPSALHNS	SATLNNSMELDNNTLTKMTQITTC-----	EDISMVHESI AVELNSN 623	
Qy	184	ALSNSFTKVI	QNNQESFVG---TF-----	TNANVQPSNYSFVA--FSADVPVNY 229	
Db	624	KEQPFQGD	ETLQKNDTNTGANTFQGHNETSQILMNVDSEAVNTSKISTYSAEFLSINQ	683	
Qy	230	KYA--RRTV	YNGDEPSSKILANTNSITDVSIVYSLAGTN-----	TKYQFSFSGPSGYL 283	
Db	684	SISKRRSLNS	ARESPPRAVLENS-----	IMSNQGTWEALTEYR---QNKTWQTSQD 734	
Qy	284	YFPYKLVKAAD	ANNVGLQYKLNNGNVQOQEFATSTANNTANPTPAVDKIAKIVLSG	343	
Db	735	SMP-----	SMSLNDGSRDILAMNTSVR-----	SPHLNSKTAAPQPSL-----	773

	Query Match	6.28;	Score	124.5;	DB	5;	Length	1225;
	Best Local Similarity	21.68;	Pred. No.	9.1;				
	Matches	98;	Conservative	51;	Mismatches	155;	Indels	149; Gaps
Qy	51	SSAYSEAEYNNN-----LNATLEQLKMKARTNLESAINQANTDKTTFDNEHP- : :	103	NLVEAY				
Dd	358	NAALNEAQFAKNSCSYTHITGYDLGDFALANNLSKLTA---TDENHFFATPANLA : :	413	AAP				
Qy	104	KALKATTLEQRATNLEGLSSRAYNOIRNLVDLYNKASLTIKTKLDPLNG : :	152					
Dd	414	DNIAQTINIGIQREVTFVAGPFIIVMLTO-----SGDVTHLLNVNSGVHYDVSTFKL : :	468					
Qy	153	-----GTLLDSNEITA-----NKNIN-NWTLSTINEQKTNADALSNSFIIKKVIQN : :	196					
Dd	469	TWTGTTLSSSEAITIYRIYADLDYIQNDIPVNTTGAIGFDLGGFDT-----N : :	517					
Qy	197	NEQSFEVGFTFNANYPQSNYSFVAFSADVTPNYKYARRTVMNGDEPSRILANTSIDV : :	256					

Db 518 TEAKL--TYTNSGEPNQ--LIFPRPTVKGIVKRRHYLVNKGQPIQANGTVVSSL 573
QY 257 S-----WI-----YSLAGTNT-----KYQF----- 271
Db 574 SEAHVLSQDFFLPSCGGGHIVPKWIKDKTTEALQIYVSPPTNTVITADGKRYREVEVP 633
QY 272 -SFSNYG-----PSTGYLYFPYKLVK-----ADANNVGLQYKLNNGVQ 310
Db 634 GSTPNPGQIGISWKKPAGNAYFAYKLLNVMGTTDQSEWDVTSNWTGAQVPLTG---E 690
QY 311 QVEFATSTANNTTAPVADBIKV-----AKIVLSGLRFGQNTIELSVPTG-----EGN 361
Db 691 DVEFATTENFG-----SPAVADLHVPTTNPKEII--GNLNNSDKDLVVTSSQLTNGV 742
QY 362 MNKVAPMIGNIYLSNENNAKDI-----PGYRR 389
Db 743 VEDNNPNVGTIVVKSSKDNPTGTLLFANPGYNQ 775

RESULT 7

US-09-581-286A-484
; Sequence 484, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHEN, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 484
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (983)..(983)
; OTHER INFORMATION: xaa can be any naturally occurring amino acid
US-09-581-286A-484

Query Match

Best Local Similarity 6.2%; Score 124.5; DB 5; Length 1226;

Matches 98; Conservative 51; Mismatches 155; Indels 149; Gaps 22;

QY 51 SSAYSEAEVTNNN-----LNATLQKMAKTNLESAINQANTDKTTFDNEHP-NLVEAY 103

Db 359 NAAINEAQFAKNSGYTIHTIGYDLGDFALANNSKLTA-----TDENHFFATPANLAAAF 414
QY 104 KALKTTLEQRATNLEGLSSTAYNQIRNNLVNLYNKASSLITKTLDPLNG----- 152
Db 415 DNIAQTINIGIQGEVTDVAPGFIVKNTQ-----SGDVTHLLNVSNGTVHYDVSTKKL 469
QY 153 -----GTLDSNETTA-----NKNIN-NTLSTNEQKTNADALSNSFIKKVIQON 196
Db 470 TWTGTILSSSEATITIRIYADLDYIQNNDIPVNTTSAICPDLGEDT-----N 518
QY 197 NEOSFVGTITNANVQSNYSFVAFSDVPTVNTKYARRTVWNGDESSRILANTNITDV 256
Db 519 TEAKL--TYTNSGEPNQ--LIFPRPTVKGIVKRRHYLVNKGQPIQANGTVVSSL 574
QY 257 S-----WI-----YSLAGTNT-----KYQF----- 271
Db 575 SEAHVLSQDFFLPSCGGGHIVPKWIKDKTTEALQIYVSPPTNTVITADGKRYREVEVP 634
QY 272 -SFSNYG-----PSTGYLYFPYKLVK-----ADANNVGLQYKLNNGVQ 310
Db 635 GSTPNPGQIGISWKKPAGNAYFAYKLLNVMGTTDQSEWDVTSNWTGAQVPLTG---E 691
QY 311 QVEFATSTANNTTAPVADBIKV-----AKIVLSGLRFGQNTIELSVPTG-----EGN 361
Db 692 DVEFATTENFG-----SPAVADLHVPTTNPKEII--GNLNNSDKDLVVTSSQLTNGV 743
QY 362 MNKVAPMIGNIYLSNENNAKDI-----PGYRR 389
Db 744 VEDNNPNVGTIVVKSSKDNPTGTLLFANPGYNQ 776

RESULT 8

US-09-581-286A-351
; Sequence 351, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHEN, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 351
; LENGTH: 1251

Db 412 KLDVTDQFDVYISNDKATVDMKGTSSNQYIIQQVAYPDNSSTDNGKIDYITLDDT 471
QY 266 NTKYQF--SPSNY-GPSTG-----YLYFPYKLVKAADANNVGLQ---YKLNNGN 308
Db 472 KTKYSNSYSNYSNGSSSTANGDKYKNGLDYVWEDTKDQKQDANEKIGKGVVILKDSN 531
QY 309 VQOVEFATSPSANTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 358
Db 532 GKELD-RTTYDEN-----GKYQFTGLSNGTYSVEFSTPAG 565

RESULT 11

PCT-US03-20460-14
; Sequence 14, Application PC/TUS0320460
; GENERAL INFORMATION:
; APPLICANT: Iowa State University Research Foundation, Inc. and Department of
; APPLICANT: Agriculture for
; APPLICANT: and on behalf of the State of New South Wales
; TITLE OF INVENTION: Immunogenic Mycoplasma Hypopneumoniae
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 08411/035W01
; CURRENT APPLICATION NUMBER: PCT/US03/20460
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392,632
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 945
; TYPE: PRT
; ORGANISM: Mycoplasma hypopneumoniae
PCT-US03-20460-14

Query Match 5.9%; Score 118; DB 1; Length 945;
Best Local Similarity 21.3%; Pred. No. 16;
Matches 84; Conservative 46; Mismatches 129; Indels 136; Gaps 16;
QY 33 KAMTLASLDQYAKIEASLSAYSEAEVNN--NLNATLEQLKMAKTNLESAINQANTDKT 90
Db 441 KEFKMQEYEDRAKL-----STREIKETIDKLANLAQVSNLSPSDEVRAVYLLNTGKY 495
QY 91 TFDNEHPNLVEAYKALKLTLEQKATNLEGLSSTAYNQIRNLDVLYNKASSLITKTLDP 150
Db 496 LFDDE-----IQEKTNLKIIIEQA--RMKADTKNLAPKVPSPLOK-----534
QY 151 NGTLLDSNEITANKNINNTLSTINEQKTNADALSNSIKKVIQNEQSFVGTFTNANV 210
Db 535 -----PTTSATSSGTTKTSTGTGKVKSVSAFSDIISMK--NQPEQ-----TTKNGQV 579
QY 211 QPSNYSFVAFSADVTPVNYKARTVWNGDEPSSRLANT--NSIT-----D 255
Db 580 QASSTS-----QSPKSSLSQNSQNSITLEKFGHTIWKLLN 616
QY 256 VSMYSLAGTNTKYQFSFGPSTGYLYFPYKLVKAAD-----ANNYGLQYKLNNG-- 307
Db 617 TSQIYNFENTQGGYTISIED-----DKLVDFDKLVSKADRAIIYQGSKISLGLINSRKS 671
QY 308 -----NQOVEFATSTSAN-----NTTANPTPAVD 332
Db 672 AYDEIKQFSPDLFDATIGEQSDYKKNKKQDKTLKSLRDLGMNGFGVYKPEKTSNPOENVL 731
QY 333 EIKVA---KIVLSGLRFGQNTIELSVPTGEGNNK 364
Db 732 KLOTGSEQKPLPLGLSLGLIYTAFTV-----NNINK 762

RESULT 12

PCT-US03-27401-316
; Sequence 316, Application PC/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND

; TITLE OF INVENTION: PREVENTION OF ACTIVE INFECTION
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-316

Query Match 5.8%; Score 116.5; DB 1; Length 1881;
Best Local Similarity 15.9%; Pred. No. 33;
Matches 74; Conservative 89; Mismatches 162; Indels 139; Gaps 17;
QY 9 DANPNNGQTOLEAARMELTDLINAKAWTLASLDY-----AKIEASLS 52
Db 441 NSETNKLKTAIDALNVDKTELNNTIADAKTKVKEHYSDRSWQNLQTEVTKAEKVAANTDA 500
QY 53 AYSEAEVNNNLNATLEQLKMAKTNLESAINQANTDKT-----TFDNEHPNLVEA 102
Db 501 KQSEVNEAVEKLTATIE--KLVELSEKPILTLTSTDKILEREAVAKYLTLENQNKTKIKS 558
QY 103 Y-----KALKTTLEQKATNLEGLSSTAYNQIRNLDVLYNKASSLI 143
Db 559 ITAELKKGEVINTVVTDDKVTTETISAFAKNLE-----YKEYTLSTTMIYDRNGEE 613
QY 144 TKTLDPNGTLLDSNEITANKNINNTLSTINEQKTNADALS-----NSFKKVIQNN 197
Db 614 TETLE--NQIQDLKLLKVELKNIKRTDLIKYENGKETNESLITIPDDKSNYYLKITSNN 671
QY 198 EQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKARTVWNGDEPSSRLANTNSITDVS 257
Db 672 QKT-----TLLAVANIETTVNG--TPYKVTAIADNLVSR 706
QY 258 WIYSLAGTNTKYQFSFN--GPST--GYLYFPYKLVKAADANNVGLQYKLNNGNVQOVE 313
Db 707 -----ADNKFEEVHYIEKPKVHDNYYVNFKELVEAIQNDPSKEYRLGQ-----752
QY 314 FATSTSANNTANPTP-----AYDEIK-----VAKIVLSGLRFGQN 349
Db 753 ---SMGARNVVPNGKSYIETKFTGKLLSSEKQFAITELEHPLFNVTINATINNANFV 809
QY 350 TIELSVPTGEGNNKVA-----PMICNTIYVLSNENNADKIPCY 387
Db 810 ETERS---GODNTASLANTMKGSSVITNVKITGTLSGRNNVAGF 850

RESULT 13

US-10-650-274-142
; Sequence 142, Application US/10650274
; GENERAL INFORMATION:
; APPLICANT: GLENN MATTHEW
; APPLICANT: HAVUKKALA, ILKKA J
; APPLICANT: LOBBERS, MARK WILLIAM
; APPLICANT: DEKKER, JAMES
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,
; TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING
; FILE REFERENCE: 11000.1073
; CURRENT APPLICATION NUMBER: US/10/650,274
; CURRENT FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 142
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: LACTOBACILLUS RHAMNOSUS
US-10-650-274-142

Query Match 5.7%; Score 115; DB 6; Length 1222;
Best Local Similarity 24.6%; Pred. No. 28;
Matches 47; Conservative 33; Mismatches 87; Indels 24; Gaps 5;

QY 18 QLEARMELTDLINAKMTLASLDYAKIEASLSAYSEAEVNNLNATLEOLKMAKTN 77
DB 823 QLOKAKTTLDLLIQAKAKATSQAQYDATTNALNTAIGSAQTALAKADATIELTTATTQ 882
QY 78 LESAINQANTOKTFD-----NEHPNLVEAY-----KALKTTLEQATWLE 118
DB 883 LTTAVNQL-VDKLPADQQAALLNKIQSAKEAFGTDLGQOTDPSTCKTLNAELDAVAQAQT 941
QY 119 GLSTAYNOIRNNLVLYNKASSLITKTL---DPLNGGTLSDSNEITANKNNINLTSTI 175
DB 942 AGTSTA-DOETENFKNVLDALNQLAKTIKAATPKVGVGNANDTTGKTWYGDVAVIAAG 1000
QY 176 NEQKTADALS 186
DB 1001 TAAKTDTEKIA 1011

RESULT 14
US-10-047-676B-7
; Sequence 7, Application US/10047676B
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; APPLICANT: Caulfield, Page W.
; APPLICANT: Chen, Ping
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17403/22
; CURRENT APPLICATION NUMBER: US/10/047,676B
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 09/627,376
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 990
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-047-676B-7

Query Match 5.7%; Score 114.5; DB 6; Length 990;
Best Local Similarity 21.0%; Pred. No. 25;
Matches 90; Conservative 62; Mismatches 154; Indels 123; Gaps 19;

QY 9 DANPNNGQTQLEARMELTDLINAKM-----TLASLDYAKIEASLS 52
DB 274 DRNIGEGNSILALENKMQUIVKANAYLRVLDYDHAELKLAQHTKSSQLNLIKVLSPSS 333
QY 53 AYSEAEVNNN-----LNATLEOLKMAKTNLESAINQANTOKTFD 93
DB 334 AVNSQKEIKNYHEKFIARYGYEQLVPLQLLNST-SGLGFPKGYSTQEVSKONNE----D 388
QY 94 NEHPNLVEAY-----KALKTTLE--QRATNLEGLSSTAYNOIRNNLVLYN-KASSLITK 145
DB 389 SKNOKIIEFLQKPEKALRDKGEIILSDDDLDKLNFDTEQOISGELYCFYFNPKSKKLEVS 448
QY 146 TLDPLNGGTLSDSNEITANKNNINLTSTINEQKTN---ADALSNSFKKVIQNNEQSFV 202
DB 449 SL----GVSQMLGNTFGFHSKLPNTIVTKVNKTKEIFTEAYPNTII---TQLEVPYF 501
QY 203 GTFTNANYQPS-----NYS-----FV-AFSADVTPVNYKIARTVW----- 237
DB 502 GRGGNIMISNLKSHOLELRNYTTKKEMSINDIYVRATSEELYFYSKKYKRVIFVNMNM 561
QY 238 ----NGDEPSSRLANTNSIDVSWIYSLAGTNTKYQSFNSGPGST-----GYLYFPYK 288
DB 562 FNYINGSKLLRFLLEVSN-----DFQNTITPITLGLSDSYNHNVP 602
QY 289 LVKAADANNVQYKLNNGNQOQVEFATSTANN--TTANPTPAVDKIAKIVLSGLRF 346
DB 603 IYK-----DIIRKPTWIRKSEAKTLDLSKLNLTNNVPPFVPMKYTDQIIYLDL--- 653

QY 347 QNTIELSV 355
DB 654 -SRTIDLTM 661

RESULT 15
US-10-425-114A-51665
; Sequence 51665, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51665
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700221631_FLI.pap
US-10-425-114A-51665

Query Match 5.7%; Score 114; DB 6; Length 647;
Best Local Similarity 25.3%; Pred. No. 18;
Matches 47; Conservative 33; Mismatches 72; Indels 34; Gaps 7;

QY 18 QLEARMELTDLINAKMTLASLDYAKIEASLSAYSEAEVNNLNATLEOLKMAKTN 77
DB 82 QLEKARYELANAIKAKNEALSQVDDAVRA-----SEAKT-----QVEVHLIAEVTH 127
QY 78 LESAIN---QANTDKTTFDNEHPNLVEAYKALKTTL-----EQATNLEGLSSTAYNOI 128
DB 128 LKGLVDSKVQDGNRKTA--ERIQNLEENFALKLEQKKAAEKAIRLERMIDELOSDV 185
QY 129 ---RN-----NLVDLYNKASSLITKTLDPNGGTLSDSNEITANKNNINLTSTINEQK 179
DB 186 DDARFSGSKOLADEWQKQALLVRLVEEADQSNILKESUNSAWEELDSTSLRDRE 245
QY 180 TNADAL 185
DB 246 SEVAAL 251

Search completed: October 8, 2003, 17:40:26
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:29:36 ; Search time 394 Seconds
(without alignments)
909.919 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSITKDKANPNNGQTGLE.....SSNENNADKIPGYRRPGTFL 394

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
- 32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2004	100.0	456	15	US-09-147-052-2 Sequence 2, Appl

2	2004	100.0	1086	15	US-09-147-052-4	Sequence 4, Appl
3	1997	99.7	456	24	US-09-901-572A-3	Sequence 3, Appl
4	1929	96.3	384	27	US-10-131-591A-12	Sequence 12, Appl
5	1777	88.7	357	27	US-10-131-591A-23	Sequence 23, Appl
6	746	37.2	645	22	US-09-791-537-105581	Sequence 105581
7	744	37.1	650	22	US-09-791-537-116408	Sequence 116408
8	713	35.6	647	4	US-08-014-631-7	Sequence 7, Appl
9	713	35.6	647	6	US-08-230-312-7	Sequence 7, Appl
10	713	35.6	647	22	US-09-791-537-7385	Sequence 7385, Ap
11	713	35.6	647	22	US-09-791-537-34555	Sequence 34555, A
12	176.5	8.8	1291	32	US-60-257-931-3393	Sequence 3393, Ap
13	176.5	8.8	1291	32	US-60-257-931-3393	Sequence 3393, Ap
14	176.5	8.8	2398	28	US-10-282-122A-70176	Sequence 4114, A
15	176.5	8.8	6281	1	PCT-US02-03987-12996	Sequence 70176, A
16	176.5	8.8	6281	23	US-09-815-242-12996	Sequence 12996, A
17	176.5	8.8	6281	26	US-10-072-851-12996	Sequence 12996, A
18	176	8.8	1095	21	US-09-710-279-3154	Sequence 3154, Ap
19	175	8.7	1947	12	US-08-827-356-5566	Sequence 5566, Ap
20	175	8.7	1947	20	US-09-611-529-4400	Sequence 4400, Ap
21	175	8.7	1947	25	US-09-950-084-4200	Sequence 4200, Ap
22	175	8.7	3351	12	US-08-827-356-5564	Sequence 5564, Ap
23	175	8.7	3351	20	US-09-611-529-4562	Sequence 4562, Ap
24	175	8.7	3351	25	US-09-950-084-4562	Sequence 4562, Ap
25	175	8.7	6641	28	US-10-282-122A-70580	Sequence 70580, A
26	175	8.7	6713	28	US-10-282-122A-43811	Sequence 43811, A
27	175	8.7	10182	26	US-10-092-411A-3159	Sequence 3159, Ap
28	175	8.7	10203	18	US-09-450-969-4098	Sequence 4098, Ap
29	171	8.5	2086	1	PCT-US02-03987-5639	Sequence 5639, Ap
30	171	8.5	2086	23	US-09-815-242-5639	Sequence 5639, Ap
31	171	8.5	2086	26	US-10-072-851-5639	Sequence 5639, Ap
32	171	8.5	2086	32	US-60-242-578-948	Sequence 948, App
33	171	8.5	2086	32	US-60-253-625-2292	Sequence 2292, Ap
34	171	8.5	2086	32	US-60-257-931-3208	Sequence 3208, Ap
35	171	8.5	2086	32	US-60-269-308-4231	Sequence 4231, Ap
36	171	8.5	5795	23	PCT-US02-03987-12610	Sequence 12610, A
37	171	8.5	5795	23	US-09-815-242-12610	Sequence 12610, A
38	171	8.5	5795	26	US-10-072-851-12610	Sequence 12610, A
39	164.5	8.2	2481	28	US-10-282-122A-43762	Sequence 43762, A
40	164.5	8.2	2495	12	US-08-827-356-3899	Sequence 3899, Ap
41	164.5	8.2	2495	20	US-09-611-529-6143	Sequence 6143, Ap
42	164.5	8.2	2495	25	US-09-950-084-6143	Sequence 6143, Ap
43	162.5	8.1	2659	29	US-10-311-879-28	Sequence 28, Appl
44	162	8.1	1029	1	PCT-US02-03987-5885	Sequence 5885, Ap
45	162	8.1	1029	23	US-09-815-242-5885	Sequence 5885, Ap

ALIGNMENTS

RESULT 1

US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

RECOMBINANT VECTOR,

Query Match 100.0%; Score 2004; DB 15; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.5e-161;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 60
DB 63 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 122

QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 120
DB 123 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 182

QY 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 180
DB 183 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 242

QY 181 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 240
DB 243 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 302

QY 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 300
DB 303 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 362

QY 301 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
DB 363 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422

QY 361 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 394
DB 423 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 456

RESULT 2
US-09-147-052-4
; Sequence 4, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid

US-09-147-052-4

Query Match 100.0%; Score 2004; DB 15; Length 1086;
Best Local Similarity 100.0%; Pred. No. 5.8e-161;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 60
DB 693 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 752

QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 120
DB 753 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 812

QY 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 180
DB 813 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 872

QY 181 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 240
DB 873 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 932

QY 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 300
DB 933 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 992

QY 301 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
DB 993 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 1052

QY 361 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 394
DB 1053 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 1086

RESULT 3
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S

US-09-901-572A-3

Query Match 99.7%; Score 1997; DB 24; Length 456;
Best Local Similarity 99.5%; Pred. No. 6.1e-161;
Matches 392; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 60
DB 63 GCMSTTKDANPNNGQTOLEAARMELTDLINAKAMTASLDQYAKIEASLSAYSEATV 122

QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 120
DB 123 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 182

QY 121 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 180
DB 183 SSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKT 242

QY 181 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 240
DB 243 NADALSNSFKKVIQNNESFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTVWNGD 302

QY 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 300
DB 303 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFFPKYKLVKAADANNVGL 362

QY 301 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
DB 363 QYKLNNGNVQVQEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422

QY 361 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 394
DB 423 NMKVAPMIGNIYLSNENNAADKIPGYRRPGTFL 456

RESULT 4
US-10-131-591A-12
; Sequence 12, Application US/10131591A
; GENERAL INFORMATION:

```
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Query Match      96.3%; Score 1929; DB 27; Length 384;
Best Local Similarity 99.5%; Pred. No. 2.9e-15;
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  2 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTLDINAKMTLASLDYAKIEASLSAYSEAEVTN 61
Db  3 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTLDINAKMTLASLDYAKIEASLSAYSEAEVTN 62
QY  62 NNLNATLEQKMAKTNLESAINQANTDKTTDFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db  63 NNLNATLEQKMAKTNLESAINQANTDKTTDFDNEHPNLVEAYKALKTTLEQRATNLEGLS 122
QY  122 STAYNQIRNLDVLYNKASSLITKTLDPLNGGTLDSNEITANKNINNTLSTINEQKTN 181
Db  123 STAYNQIRNLDVLYNKASSLITKTLDPLNGGTLDSNEITANKNINNTLSTINEQKTN 182
QY  182 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
Db  183 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 242
QY  242 PSSRLANTNSITDVSIIYSLAGNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 301
Db  243 PSSRLANTNSITDVSIIYSLAGNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 302
QY  302 YKLNGNVOQVEFATSTSANNTANPTPAVDKIVLKGRLFGQNTIELSVPTGEGN 361
Db  303 YKLNGNVOQVEFATSTSANNTANPTPAVDKIVLKGRLFGQNTIELSVPTGEGN 362
QY  362 MNKVAPMIGNIYLSNENNAKD 383
Db  363 MNKVAPMIGNIYLSNENNAKD 384

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified Tti-1 portion (downstream of BglI) of
; OTHER INFORMATION: pNZ40K-S
US-10-131-591A-23

Query Match      88.7%; Score 1777; DB 27; Length 357;
Best Local Similarity 98.6%; Pred. No. 2.4e-142;
Matches 352; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  27 TDLINAKAMTLDYAKIEASLSAYSEAEVTNNLNATLEQKMAKTNLESAINQAN 86
```

```
Db  1 TDLINAKAMTLDYAKIEASLSAYSEAEVTNNLNATLEQKMAKTNLESAINQAN 60
QY  87 TDKTTDFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNOIRNLDVLYNKASSLITKT 146
Db  61 TDKTTDFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNOIRNLDVLYNKASSLITKT 120
QY  147 LDPLNGGTLDSNEITANKNINNTLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFT 206
Db  121 LDPLNGGTLDSNEITANKNINNTLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFT 180
QY  207 NANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSIIYSLAGTN 266
Db  181 NANVQPSOYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSIIYSLAGTN 240
QY  267 TKYQFSNYPSTGYLYFPYKLVKAADANNVGLQYKLNGNVOQVEFATSTSANNTAN 326
Db  241 TKYQFSNYPSTGYLYFPYKLVKAADANNVGLQYKLNGNVOQVEFATSTSAOQTAN 300
QY  327 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKVAPMIGNIYLSNENNAKD 383
Db  301 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKVAPMIGNIYLSNENNAKD 357

RESULT 6
US-09-791-537-105581
; Sequence 105581, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105581
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
US-09-791-537-105581

Query Match      37.2%; Score 746; DB 22; Length 645;
Best Local Similarity 42.3%; Pred. No. 4.8e-54;
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;

QY  9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLDYAKIEASLSAYSEAEVTN 61
Db  44 DTNPGDGGGMNAASQELAAARMGLTTFVDSKAKNLGLYVDYKKTQDTLTIKAYDAKTVL 103
QY  62 NNLNATLEQKMAKTNLESAINQANTDKTTDFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db  104 DNSSSTONLEAKTRLETAINTAATSKQTFDEQHAELVKVYEELKTLTSLNETATLAPYA 163
QY  122 STAYNQIRNLDVLYNKASSLITKTLDPLNGGTLDSNEITANKNINNTL--STINEQK 179
Db  164 AAQYAGIKMHLGSLYDAGAKITKTLEPVEGDP-LTADVVMKNTKIIEAKIDEVLNPKR 222
QY  180 TNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPV-----NKY 231
Db  223 ENATKLADSFVKQVLVKEKITGVEEAHN-KAOPANYSFVGSVDITGTGTGTSIPNDY 281
QY  232 ARRTVW-NGDEPSSRLANT-----NSITDVSIIYSLAGTNKYQFSNYPSTGY 282
Db  282 AQRITFTNGDEP--RSISNTPADGOTMVOPLSNVSIYSLAGTGAKTYLFTFYGPSTGY 339
QY  283 LYFPYKLVKAADANNVGLQYKLNGNVOQVEFATSTSA-----NNTANPTPAVDKIVA 337
Db  340 LYFPYKLVNTSDMKLGLYKLDN-----ATEPSAITFGSEOTMGKTFVTVDINVA 391
QY  338 KIVLSGLRFGQNTIELSVPTGEGNMKVAPMIGNIYLSNENNAKD 384
```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,312
FILING DATE: April 20, 1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: DIGIGLIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 88012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 647 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-230-312-7

Query Match 35.6%; Score 713; DB 6; Length 647;
Best Local Similarity 41.1%; Pred. No. 3.2e-51;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;
QY 9 DANPNNGQ-----TOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSSETV 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTKAYDAKTVL 107
QY 62 NNLNATLEQLMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGTGLDSNEITANKNINNTL--STINEOK 179
DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIIVEAIDKDEVLNPK 226
QY 180 TN-----ADALNSFTKKVQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 227
DB 227 ENATKLADSLSSIVKKTGVEE-----AHNKAQAPANYSFVGYKRWYTELLLDKQVFP 279
QY 228 NYKVARRTVM-NGDEPSSRIILANT-----NSITDYSWIYSLAGTWTKYQFSNSTYGP 278
DB 280 NWDYAQRITFTNSDEP--RSISNTPADGQTMAQPLSNVSWIYSLAGTGAKYTLEFTYYPG 337
QY 279 STGYLYFPYKLVKAADANNVGLQYKLNNGNQVQEFATSTSA-----NNTTANPTPAYDE 333
DB 338 STGYLYFPYKLVNTSDQVKLGLEYKLN-----ATKPSAITFGSDQTMNGKTPTYND 389
QY 334 IKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYSSNENNADKI 384
DB 390 INVAKVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNNKI 435

RESULT 10

US-09-791-537-7385
Query Match 35.6%; Score 713; DB 22; Length 647;
Best Local Similarity 41.1%; Pred. No. 3.2e-51;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;
QY 9 DANPNNGQ-----TOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSSETV 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTKAYDAKTVL 107
QY 62 NNLNATLEQLMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGTGLDSNEITANKNINNTL--STINEOK 179
DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIIVEAIDKDEVLNPK 226
QY 180 TN-----ADALNSFTKKVQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 227

ORGANISM: Mycoplasma gallisepticum
US-09-791-537-7385
Query Match 35.6%; Score 713; DB 22; Length 647;
Best Local Similarity 41.1%; Pred. No. 3.2e-51;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;
QY 9 DANPNNGQ-----TOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSSETV 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTKAYDAKTVL 107
QY 62 NNLNATLEQLMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGTGLDSNEITANKNINNTL--STINEOK 179
DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIIVEAIDKDEVLNPK 226
QY 180 TN-----ADALNSFTKKVQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 227
DB 227 ENATKLADSLSSIVKKTGVEE-----AHNKAQAPANYSFVGYKRWYTELLLDKQVFP 279
QY 228 NYKVARRTVM-NGDEPSSRIILANT-----NSITDYSWIYSLAGTWTKYQFSNSTYGP 278
DB 280 NWDYAQRITFTNSDEP--RSISNTPADGQTMAQPLSNVSWIYSLAGTGAKYTLEFTYYPG 337
QY 279 STGYLYFPYKLVKAADANNVGLQYKLNNGNQVQEFATSTSA-----NNTTANPTPAYDE 333
DB 338 STGYLYFPYKLVNTSDQVKLGLEYKLN-----ATKPSAITFGSDQTMNGKTPTYND 389
QY 334 IKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYSSNENNADKI 384
DB 390 INVAKVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNNKI 435

RESULT 11

US-09-791-537-34555
Sequence 34555, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE OF INVENTION: METHODS OF USE THEREOF
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 34555
LENGTH: 647
TYPE: PRT
ORGANISM: Mycoplasma gallisepticum
US-09-791-537-34555

Query Match 35.6%; Score 713; DB 22; Length 647;
Best Local Similarity 41.1%; Pred. No. 3.2e-51;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;
QY 9 DANPNNGQ-----TOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSSETV 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYVDYKKTQNTLTKAYDAKTVL 107
QY 62 NNLNATLEQLMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGTGLDSNEITANKNINNTL--STINEOK 179
DB 168 DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIIVEAIDKDEVLNPK 226
QY 180 TN-----ADALNSFTKKVQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 227

Db 227 ENATKADSLSSIVKAITGVEE-----AHNKAQAPANTSFVGYKRWYTELLELDKQVFP 279
Qy 228 NYKARTVW-NGDEPSSRLANT-----NSITDVSIIYSLAGTNTKYQFSFNKYP 278
Db 280 NWDYAQRTIETNSDEP--RSISNTPADQMTQAQPLSNVSIYSLAGTCAKYLEFTYYGP 337
Qy 279 STGYLYPPYKLVKAADANNVGLQKLNNGVQVFEATSTSA-----NNTANPTPAVDE 333
Db 338 STGYLYPPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPVND 389
Qy 334 IKVAKIVLSGLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSNENNAKDI 384
Db 390 INAKVTLANLNGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 435

RESULT 12

US-60-257-931-3393
; Sequence 3393, Application US/60257931
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, R
; APPLICANT: Ohlsen, KL
; APPLICANT: Zyskind, JW
; APPLICANT: Trawick, JD
; APPLICANT: Wall, D
; TITLE OF INVENTION: Genes identified as essential in *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*
; FILE REFERENCE: ELITRA.017PR4
; CURRENT APPLICATION NUMBER: US/60/257,931
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 3592
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3393
; LENGTH: 1291
; TYPE: PRT
; ORGANISM: *Staphylococcus aureus*
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1291)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-257-931-3393

Query Match 8.8%; Score 176.5; DB 32; Length 1291;
Best Local Similarity 22.3%; Pred. No. 5.1e-05;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;
Qy 13 NNGQTOLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE-----VNNLNAT 67
Db 799 SNTSTQNTAMANLQNGINDKNTLAS-ENYHDADSDKKTAYTQAVTNAENILNKGSGN 857
Qy 68 LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNO 127
Db 858 LD-----KTAVENALSQVANAKALNGNH-NLEQAKSNANTTI-----NGLQHLTAAQDK 907
Qy 128 IRNLDVLYNKAS-SLITKTLDPLNG--GTLLDSNEITTANKNINNTLSTINEQKT---- 180
Db 908 LKQOVQOQAVAGVDVTKSSANTLNGAMGTLRNSIQDTATNGQYLDATERNKTNYNN 967
Qy 181 -----NADALSNSFKKVIQNNESQFVGFTTANVQPSNYSFVAFSADVT 225
Db 968 AVDSANGVINATSNPNMDANAIQIATQVTSKNDLGDTHNLTQAKQT-----ATNAIDG 1022
Qy 226 PNYKYARTVWNGDEPSSRLANTNSITDVSIIYSLA-----GNTKYQFSF 273
Db 1023 ATNLNKAQDKALKQAQVTSQAQRVANVTSIQQTANELNTAMGQLQHGIDDENATKQTKYRD 1082
Qy 274 SNYGSTGYLYPPYKLVKAADA-----NNVGLQYKLN-----NGNQVQVEF 314
Db 1083 AEQSKKTAY-----DQAVAAAKAILNKQTSNSDKAAVDRAALQOQVTSKDALNGDAKLABA 1138
Qy 315 ATSTSAN-----NTANPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMN----- 363
Db 1139 KAAAKQNLGTLNHTNAQRTDLEGOINQATTV-----DGVNTVKTNANTLDGAMNSLOGS 1193

RESULT 14

US-10-282-122A-70176
; Sequence 70176, Application US/10282122A

Qy 364 ---KVAPMIGNIYLSNEN 379
Db 1194 INKDATLRNQNYLDADES 1212
RESULT 13
US-60-269-308-4414
; Sequence 4414, Application US/60269308
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, RW
; APPLICANT: Ohlsen, KL
; APPLICANT: Zyskind, JW
; APPLICANT: Trawick, JD
; APPLICANT: Wall, D
; TITLE OF INVENTION: Identification of Essential Genes in *Staphylococcus aureus*, *Ps*
; TITLE OF INVENTION: *aeruginosa*, *Klebsiella pneumoniae*, *Salmonella typhimurium*, an
; FILE REFERENCE: ELITRA.017PR5
; CURRENT APPLICATION NUMBER: US/60/269,308
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 4774
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4414
; LENGTH: 1291
; TYPE: PRT
; ORGANISM: *Staphylococcus aureus*
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1291)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-269-308-4414

Query Match 8.8%; Score 176.5; DB 32; Length 1291;
Best Local Similarity 22.3%; Pred. No. 5.1e-05;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;
Qy 13 NNGQTOLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE-----VNNLNAT 67
Db 799 SNTSTQNTAMANLQNGINDKNTLAS-ENYHDADSDKKTAYTQAVTNAENILNKGSGN 857
Qy 68 LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNO 127
Db 858 LD-----KTAVENALSQVANAKALNGNH-NLEQAKSNANTTI-----NGLQHLTAAQDK 907
Qy 128 IRNLDVLYNKAS-SLITKTLDPLNG--GTLLDSNEITTANKNINNTLSTINEQKT---- 180
Db 908 LKQOVQOQAVAGVDVTKSSANTLNGAMGTLRNSIQDTATNGQYLDATERNKTNYNN 967
Qy 181 -----NADALSNSFKKVIQNNESQFVGFTTANVQPSNYSFVAFSADVT 225
Db 968 AVDSANGVINATSNPNMDANAIQIATQVTSKNDLGDTHNLTQAKQT-----ATNAIDG 1022
Qy 226 PNYKYARTVWNGDEPSSRLANTNSITDVSIIYSLA-----GNTKYQFSF 273
Db 1023 ATNLNKAQDKALKQAQVTSQAQRVANVTSIQQTANELNTAMGQLQHGIDDENATKQTKYRD 1082
Qy 274 SNYGSTGYLYPPYKLVKAADA-----NNVGLQYKLN-----NGNQVQVEF 314
Db 1083 AEQSKKTAY-----DQAVAAAKAILNKQTSNSDKAAVDRAALQOQVTSKDALNGDAKLABA 1138
Qy 315 ATSTSAN-----NTANPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMN----- 363
Db 1139 KAAAKQNLGTLNHTNAQRTDLEGOINQATTV-----DGVNTVKTNANTLDGAMNSLOGS 1193
Qy 364 ---KVAPMIGNIYLSNEN 379
Db 1194 INKDATLRNQNYLDADES 1212

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 70176
LENGTH: 2398
TYPE: PRT
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (311)..(311)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (313)..(313)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (327)..(327)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (328)..(328)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (336)..(336)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (350)..(350)
OTHER INFORMATION: X-any amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (362)..(362)
OTHER INFORMATION: X-any amino acid
US-10-282-122A-70176

Query Match 8.8%; Score 176.5; DB 28; Length 2398;
Best Local Similarity 22.3%; Pred. No. 0.00013;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;
QY 13 NNGOTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAE-----VNNLNAT 67
DB 1649 SNTSTQLTAMANLQNGINDKNTNTILAS-ENYHDADSKTKTAYTQAVTNAENILNKNSGSN 1707
QY 68 LEQLKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGLSSTAYNQ 127
DB 1708 LD-----KTAVENALSQVANAKGALNGNH-NLEQAKSNANTTI-----NGLQHLTAAQDK 1757
QY 128 IRNNLVLYNKAS-SLIKTLDPLNG--GTLDSNEITANKNINNTLSTINEQKT---- 180
DB 1758 LKQOQQAAQNAGVDTVKSSANTLNGAMGLTANSIQDNTATKNGQNYLDATERNKNTYNN 1817
QY 181 -----NADALSNSFKKVIQNEQSFVGTFTNANVPQSNYSFVASDVT 225
DB 1818 AVDSANGVINATSNPNMDANAINQIATQVTSKNAIDSTHNLTAQKT-----ATNAIDG 1872
QY 226 PNYKYARRTVMNGDEPSSRILANTNSITDVSIIYSLA-----GNTKYQFSF 273
DB 1873 ATNLKAKQDKALKAQVTSQORVANVTSIQQTANELNTANGQLQHGIDDENATKQTKYRD 1932
QY 274 SNYGPSTGYLPYKLVKAADA-----NNYGLQYKLN-----NGNVQOQVEF 314
DB 1933 AEQSKKTAY----DQAVAAAKAILNKQTGSNSDKAAVDRLAQVTSKDALNGDAKLAEA 1988
QY 315 ATSTSAN-----NTANPTPAVDIEIKVAKIVLSGLRFGONTIELSVPTGEGNMN----- 363
DB 1989 KAAAKONLGTUNHITNAQRTDLEGINQATTV-----DGVNTVKTNANTLDGAMNSLOGS 2043
QY 364 ---KVAPMIGNIYLSNEN 379
DB 2044 INDKDATLRNQNYLDADES 2062

RESULT 15

PCT-US02-03987-12996
Sequence 12996, Application PC/TUS0203987
GENERAL INFORMATION:
APPLICANT: Elittra Pharmaceuticals, Inc.
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
FILE REFERENCE: ELITRA.028VPC
CURRENT APPLICATION NUMBER: PCT/US02/03987
CURRENT FILING DATE: 2002-02-02
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281
TYPE: PRT
ORGANISM: Staphylococcus aureus
PCT-US02-03987-12996

Query Match 8.8%; Score 176.5; DB 1; Length 6281;
Best Local Similarity 22.3%; Pred. No. 0.00058;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;
QY 13 NNGOTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAE-----VNNLNAT 67
DB 266 SNTSTQLTAMANLQNGINDKNTNTILAS-ENYHDADSKTKTAYTQAVTNAENILNKNSGSN 324
QY 68 LEQLKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGLSSTAYNQ 127
DB 325 LD-----KTAVENALSQVANAKGALNGNH-NLEQAKSNANTTI-----NGLQHLTAAQDK 374
QY 128 IRNNLVLYNKAS-SLIKTLDPLNG--GTLDSNEITANKNINNTLSTINEQKT---- 180
DB 375 LKQOQQAAQNAGVDTVKSSANTLNGAMGLTANSIQDNTATKNGQNYLDATERNKNTYNN 434

```

181 QY -----NADALSNSPIKKYIQNEQSFVGTFTNANVQPSNYSEVAFSADVT 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
435 Db AVDSANGVINATSNPNPDMAINOATQVSTKKNALDGTNLTQAKQI-----ATNAIDG 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226 QY PVNYKYARRTWINGDEPSSRLIANTNSITDYSWISLA-----GTNTKYQSF 273
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 Db ATNLNKAQKDALKAQVTSQARVANVTSIQQTANELNTAMGOLQHGIDDENATKQTKYRD 549
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
274 QY SNTGSPSTGYLYFPFKLYKAADA-----NNVGLQYKLN-----NGNVQQVEF 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
550 Db AEQSKTKAY----DQAVAAAKAILNKQTGNSDKAAVDRAJQQVSTKDALNGDAKLAEA 605
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 QY ATSTISAN-----NTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGSGNN----- 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
606 Db KAAKQMLGFLNHTNAQRDLDEQIINOATTV-----DGVNTVKTNANTLIDGAMNSLOGS 660
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
364 QY ---KVAPMIGNIYLSNEN 379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
661 Db INDKDATLRNQNYLDADES 679
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: October 8, 2003, 17:39:54
Job time : 396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:31:51 ; Search time 69 Seconds
(without alignments)
920.067 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTTKDANPNNGQTGLE.....SSNENADKIPGYRRPGTFL 394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2004	100.0	456	9	US-09-147-052-2
2	2004	100.0	1086	9	US-09-147-052-4
3	1997	99.7	456	12	US-09-901-572A-3
4	1929	96.3	384	15	US-10-131-591A-12
5	1777	88.7	357	15	US-10-131-591A-23
6	176.5	8.8	6281	9	US-09-815-242-12996
7	171	8.5	2086	9	US-09-815-242-5639
8	171	8.5	5795	9	US-09-815-242-12610
9	162.5	8.1	2659	12	US-10-311-879-288
10	162	8.1	1029	9	US-09-815-242-5885
11	162	8.1	1048	9	US-09-815-242-13083
12	159	7.9	496	12	US-10-311-879-29
13	159	7.9	520	15	US-10-056-052-2
14	159	7.9	936	8	US-08-781-988A-5249
15	158	7.9	2434	9	US-09-815-242-5835

16	156	7.8	807	11	US-09-820-843A-108	Sequence 108, App
17	153.5	7.7	2478	9	US-09-815-242-5816	Sequence 5816, Ap
18	153.5	7.7	2478	9	US-09-815-242-12967	Sequence 12967, A
19	145.5	7.3	2437	9	US-09-815-242-5834	Sequence 5834, Ap
20	145	7.2	857	12	US-10-032-585-7857	Sequence 7857, Ap
21	145	7.2	1215	9	US-09-815-242-5908	Sequence 5908, Ap
22	145	7.2	1269	9	US-09-815-242-13113	Sequence 13113, A
23	143.5	7.2	837	9	US-09-815-242-5883	Sequence 5883, Ap
24	143.5	7.2	875	9	US-09-815-242-13080	Sequence 13080, A
25	143	7.1	3158	9	US-09-815-242-12611	Sequence 12611, A
26	141.5	7.1	2025	9	US-09-815-242-5703	Sequence 5703, Ap
27	141.5	7.1	2368	9	US-09-815-242-5635	Sequence 5635, Ap
28	141.5	7.1	2368	9	US-09-815-242-12389	Sequence 12389, A
29	137	6.8	692	15	US-10-101-464A-897	Sequence 897, App
30	136.5	6.8	2283	12	US-10-172-502-4	Sequence 4, Appli
31	134	6.7	596	11	US-09-946-374-243	Sequence 243, App
32	134	6.7	596	12	US-10-015-387A-243	Sequence 243, App
33	134	6.7	596	12	US-10-063-735-100	Sequence 100, App
34	134	6.7	596	12	US-10-006-130A-243	Sequence 243, App
35	134	6.7	596	12	US-10-199-672-310	Sequence 310, App
36	134	6.7	596	12	US-10-006-172A-243	Sequence 243, App
37	134	6.7	596	12	US-10-187-749-310	Sequence 310, App
38	134	6.7	596	12	US-10-194-457-310	Sequence 310, App
39	134	6.7	596	12	US-10-184-642-310	Sequence 310, App
40	134	6.7	596	12	US-10-196-747-310	Sequence 310, App
41	134	6.7	596	12	US-10-015-392A-243	Sequence 243, App
42	134	6.7	596	12	US-10-017-253A-243	Sequence 243, App
43	134	6.7	596	12	US-10-173-689-310	Sequence 310, App
44	134	6.7	596	12	US-10-173-690-310	Sequence 310, App
45	134	6.7	596	12	US-10-173-691-310	Sequence 310, App

ALIGNMENTS

RESULT 1

US-09-147-052-2
; Sequence 2, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JF97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Query Match 100.0%; Score 2004; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.3e-150;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCMSTTKDANPNNGQTGLEARMETDLINAKAMTASLDQYAKTEASLSAYSEATV 60
DB 63 GCMSTTKDANPNNGQTGLEARMETDLINAKAMTASLDQYAKTEASLSAYSEATV 122
QY 61 NNNLNTATLQKMAKTNLESALNQANTDKTTFDNEHPNLVYAYKAKTTLEORATNLEGL 120
DB 123 NNNLNTATLQKMAKTNLESALNQANTDKTTFDNEHPNLVYAYKAKTTLEORATNLEGL 182
QY 121 SSTAYNQIRNNLVLDYNKASSLITTKLDPNGGTLLDSNEITANKNNINTLSTINEQKT 180

```
Db 183 SSTAYNQIRNNLDVLYNKASSLTKTLDPLNGGTLSDNEITANKNNITLSTINEQKT 242
Qy 181 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 240
Db 243 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 302
Qy 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 300
Db 303 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 362
Qy 301 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
Db 363 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
Qy 361 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 394
Db 423 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456

RESULT 2
US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; PRIOR FILING DATE: 1999-04-05
; PRIOR FILING DATE: 1996-03-29
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

Query Match 100.0%; Score 2004; DB 9; Length 1086;
Best Local Similarity 100.0%; Pred. No. 4.5e-150;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCMSTKKDANPNNGQTGLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATV 60
Db 693 GCMSTKKDANPNNGQTGLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATV 752
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
Db 753 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 812
Qy 121 SSTAYNQIRNNLDVLYNKASSLTKTLDPLNGGTLSDNEITANKNNITLSTINEQKT 180
Db 813 SSTAYNQIRNNLDVLYNKASSLTKTLDPLNGGTLSDNEITANKNNITLSTINEQKT 872
Qy 181 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 240
Db 873 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 932
Qy 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 300
Db 933 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 992
Qy 301 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
Db 993 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 1052
```

```
Qy 361 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 394
Db 1053 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 1086

RESULT 3
US-09-901-572A-3
; Sequence 3, Application US/0901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

Query Match 99.7%; Score 1997; DB 12; Length 456;
Best Local Similarity 99.5%; Pred. No. 4.7e-150;
Matches 392; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCMSTKKDANPNNGQTGLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATV 60
Db 63 GCMSTKKDANPNNGQTGLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATV 122
Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
Db 123 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 182
Qy 121 SSTAYNQIRNNLDVLYNKASSLTKTLDPLNGGTLSDNEITANKNNITLSTINEQKT 180
Db 183 SSTAYNQIRNNLDVLYNKASSLTKTLDPLNGGTLSDNEITANKNNITLSTINEQKT 242
Qy 181 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 240
Db 243 NADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGD 302
Qy 241 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 300
Db 303 EPSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNYSYPSGTYLYFPYKLVKAADANNVGL 362
Qy 301 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
Db 363 QYKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
Qy 361 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 394
Db 423 NMKVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456

RESULT 4
US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
```

```
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Query Match          96.3%; Score 1929; DB 15; Length 384;
Best Local Similarity 99.5%; Pred. No. 9a-145;
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CMSITKDDANPNNGQTGLEAARWETDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
DB 3 CMSITKDDANPNNGQTGLEAARWETDLINAKAMTSLASLDYAKIEASLSAYSEATVN 62
QY 62 NNLNATLEQKMAKNLNSAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
DB 63 NNLNATLEQKMAKNLNSAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 122
QY 122 STAYNQIRNLDVLYNKASLLTKTLDPLNGGTLLDSNEITTKANKNNNTLSTINEOKTN 181
DB 123 STAYNQIRNLDVLYNKASLLTKTLDPLNGGTLLDSNEITTKANKNNNTLSTINEOKTN 182
QY 182 ADALNSFFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
DB 183 ADALNSFFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 242
QY 242 PSRILANTNSITDYSWISYSLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 301
DB 243 PSRILANTNSITDYSWISYSLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 302
QY 302 YKLNGNVQOQVEAFSTSANNTANPTPAVDKIVAKIVLSGLRFGQNTIELSVPTGEGN 361
DB 303 YKLNGNVQOQVEAFSTSANNTANPTPAVDKIVAKIVLSGLRFGQNTIELSVPTGEGN 362
QY 362 MNKVAPMIGNIYLSNNENNAK 383
DB 363 MNKVAPMIGNIYLSNNENNAK 384

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of
; OTHER INFORMATION: pNZ40K-S
US-10-131-591A-23

Query Match          88.7%; Score 1777; DB 15; Length 357;
Best Local Similarity 98.6%; Pred. No. 8.9e-133;
Matches 352; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 27 TDLINAKAMTSLASLDYAKIEASLSAYSEATVNNNLNTATLEQKMAKNLNSAINQAN 86
DB 1 TDLINAKAMTSLASLDYAKIEASLSAYSEATVNNNLNTATLEQKMAKNLNSAINQAN 60
QY 87 TDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQIRNLDVLYNKASSLIYKT 146
DB 61 TDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQIRNLDVLYNKASSLIYKT 120
QY 147 LDPLNGGTLTDSNEITTKANKNNNTLSTINEQKNADALNSFIKKVIONNEQSFVGTFT 206
; LDPLNGGTLTDSNEITTKANKNNNTLSTINEQKNADALNSFIKKVIONNEQSFVGTFT 206

; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-815-242-12996

Query Match          8.8%; Score 176.5; DB 9; Length 6281;
Best Local Similarity 22.3%; Pred. No. 0.00032;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;

QY 13 NNGQTOLEAARMWELTDLINAKAMTSLASLDYAKIEASLSAYSEATVNNNLNTAT 67
DB 266 SNTSTOLNTAMANIQNGINDKTNTLAS-ENYHDADSKKTKTAYTQAVTNAENILNKSGSN 324
QY 68 LEQKMAKNLNSAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQ 127
DB 325 LD-----KTAVENALSOVANAKALGNH-NLEQAKSNANTYI-----NGLQHLTAAKDK 374
QY 128 IRNLDVLYNKAS-SLIYKTLDPLNG--GTLTDSNEITTKANKNNNTLSTINEQKN 180
DB 375 LKQOQVQAQNVAGVDVTKSANTLNGAMGTLRNSIQDNTATKNGQNTLDATERKNTYNN 434
QY 181 -----NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADV 225
; -----NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADV 225

; ORGANISM: Staphylococcus aureus
; US-09-815-242-12996

; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12996
```

Db 435 AVDSANGVINATSNPNMDANAIQIATQVTSKNALDGTNLHQAQOT-----ATNAIDG 489

Qy 226 PVNYKYARRVWNGDEPSSRILANTNSITDVSWIYSLA-----GTNTKYQFSF 273

Db 490 ATNLKAQKDALKAQVTSQAQRVANVTSIQQTANELNTAMGOLQHGIDENATKQTKYRD 549

Qy 274 SNTGPGSYLYFFPKLVKAADA-----NNVGLOYKLN-----NGNVQVFEF 314

Db 550 AEQSKKTAY-----DOAVAAAKAILNKOTGNSDKAAVDRLAQVTSYKDALNGDAKLAEA 605

Qy 315 ATTSAN-----NTTANPTPAVDIEKIVAKIVLSGLRFGQNTIELSVPTGEGNN-----363

Db 606 KARAKQNLGTLNHTNAQRTDLEQIQATTV-----DGVNTYKTNANTLDGAMNSLQGS 660

Qy 364 ---KVAPMIGNIYLSNEN 379

Db 661 INKDATLRNQNYLDADES 679

RESULT 7

US-09-815-242-5639

; Sequence 5639, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-05-23

; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-11-27

; PRIOR FILING DATE: 2000-11-27

; PRIOR FILING DATE: 2000-12-22

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5639

; LENGTH: 2086

; TYPE: PRF

; ORGANISM: Staphylococcus aureus

US-09-815-242-5639

Query Match 8.5%; Score 171; DB 9; Length 2086;

Best Local Similarity 22.3%; Pred. No. 0.00018;

Matches 99; Conservative 66; Mismatches 180; Indels 98; Gaps 17;

Qy 9 DANPNNGQTOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSAET-----VNNN 63

Db 916 EAVSNTG-TOLTAMANLQNGINDKANTLAS-ENYHDADSKKTAQTAVTNAENILKN 973

Qy 64 LNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALTTLEQRTNLEGLSST 123

Db 974 SGSNLD-----KAAVENALSQVTNAKALNGNH-NLEQAKSNANTT-----NGLQHLTA 1023

Qy 124 AYNQIRNLVDLYNKAS-SLITKTLDPNG--GTLDSNEITTKANKINNTLSTINEOKT 180

Db 1024 QKDKLQOVQQAQNVAGVDYFVKSSANTLNGAMGTILRNSIQDNTATNQCQNLDTATESNKT 1083

Qy 181 -----NADALNSFIKKVQIONNEQSFVGTFTNANVPQSNYSFVAFS 221

Db 1084 NYNNAVDSANGVINATSNPNMDANAIQIATQVTSKNALDGTNLHQAQOT-----ATN 1138

Qy 222 ADVTPVNYKYARRVWNGDEPSSRILANTNSITDVSWIYSLA-----GTNTKY 269

Db 1139 AIDGATNLNKAQKDALKAQVTSQAQRVANVTSIQQTANELNTAMGOLQHGIDENATKQTO 1198

Qy 270 QFSFSGYLYFFPKLVKAADA-----NNVGLOYKLN-----NGNVQ 310

Db 1199 KYRDAESKKTAY-----DOAVAAAKAILNKOTGNSDKAAVDRLAQVTSYKDALNGDAK 1254

Qy 311 QVEPATTSAN-----NTTANPTPAVDIEKIVAKIVLSGLRFGQNTIELSVPTGEGNN- 363

Db 1255 LAEAKAAAKQNLGTLNHTNAQRTALEQIQATTV-----DGVNTYKTNANTLDGAMNS 1309

Qy 364 -----KVAPMIGNIYLSNEN 379

Db 1310 LQGSINDKDATLRNQNYLDADES 1332

RESULT 8

US-09-815-242-12610

; Sequence 12610, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12610

; LENGTH: 5795

; TYPE: PRF

; ORGANISM: Staphylococcus aureus

US-09-815-242-12610

Query Match 8.5%; Score 171; DB 9; Length 5795;

Best Local Similarity 22.3%; Pred. No. 0.00077;

Matches 99; Conservative 66; Mismatches 180; Indels 98; Gaps 17;

Qy 9 DANPNNGQTOLEAARMELTDLINAKMTLASLDYAKIEASLSAYSAET-----VNNN 63

Db 3193 EAVSNTG-TOLTAMANLQNGINDKANTLAS-ENYHDADSKKTAQTAVTNAENILKN 3250

Qy 64 LNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALTTLEQRTNLEGLSST 123

```

Db      3251  SGNLD-----KAAVENALSOVTVNAGALNGNH-NLEQAKSNANTTII-----NGLQHLTTA 3300
Qy      124  AYNOIRNRLVDLYNKAS-SLITKTLDPNG--GTLDSNEIITANKNINNTLSTINEQKT 180
Db      3301  QKDKLKQOVQOQAVNAGVDVTKVSSANTLNCAMGTLRLNSIQDNTATNNGQNYLDATESNKT 3360
Qy      181  -----NADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFS 221
Db      3361  NYNNAVDSANGVTVNATSNPNNDANAINQIATQVTSKNALDGHNLTOAKQT-----ATN 3415
Qy      222  ADVTPVNYKYARRTVWNGDEPSSRIILANTNSITDVSWIYSLA-----GTNTKY 269
Db      3416  AIDGATNLNKAQDKALKAAQVTSQAQRVANVTSIQQTANELMTAMGQLQHGIDDENATKQQT 3475
Qy      270  QESFSNYGPGSTGYLYPPYKLVKAADA-----NNVCLQYKLN-----NGNVQ 310
Db      3476  KYRDAEQSKTAY-----DQVAAKAILNKQTSNSDKAAVDRALQOVTSKDALNGDAK 3531
Qy      311  QVEFATSTSAN-----NTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMN- 363
Db      3532  LAEAKAAKQNLGTLNHNITNAQRTALEGQINQATTV-----DGVNTVKTNTANTLDGAMNS 3586
Qy      364  -----KVAPMIGNIYLSNEN 379
Db      3587  LQGSINDKDATLRNQNYLDADES 3609

RESULT 9
US-10-311-879-28
; Sequence 28, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: toxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 28
; LENGTH: 2659
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-311-879-28

```

```

QY      347 GONTIELSV 355
      II : I : :
Db      2226 AQNAVEAAL 2234

RESULT 10
US-09-815-242-5885
; Sequence 5885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5885
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5885

```

	QY	347 GONTIELSV 355 : :
	Db	381 AQNAVEAAL 389 :
RESULT 11		
US-09-815-242-13083		
:	:	: Sequence I3083, Application US/09815242
:	:	: Patent No. US2002061569A1
:	:	: GENERAL INFORMATION:
:	:	: APPLICANT: Haselbeck, Robert
:	:	: APPLICANT: Ohlsen, Karl L.
:	:	: APPLICANT: Zyskind, Judith W.
:	:	: APPLICANT: Wall, Daniel
:	:	: APPLICANT: Trawick, John D.
:	:	: APPLICANT: Carr, Grant J.
:	:	: APPLICANT: Yamamoto, Robert T.
:	:	: APPLICANT: Xu, H. Howard
:	:	: TITLE OF INVENTION: Identification of Essential Genes in
:	:	: PROKARYOTES
:	:	: FILE REFERENCE: ELITRA.011A
:	:	: CURRENT APPLICATION NUMBER: US/09/815,242
:	:	: CURRENT FILING DATE: 2001-03-21
:	:	: PRIOR APPLICATION NUMBER: 60/191,078
:	:	: PRIOR FILING DATE: 2000-03-21
:	:	: PRIOR APPLICATION NUMBER: 60/206,848
:	:	: PRIOR FILING DATE: 2000-05-23
:	:	: PRIOR APPLICATION NUMBER: 60/207,727
:	:	: PRIOR FILING DATE: 2000-05-26
:	:	: PRIOR APPLICATION NUMBER: 60/242,578
:	:	: PRIOR FILING DATE: 2000-10-23
:	:	: PRIOR APPLICATION NUMBER: 60/253,625
:	:	: PRIOR FILING DATE: 2000-11-27
:	:	: PRIOR APPLICATION NUMBER: 60/257,931
:	:	: PRIOR FILING DATE: 2000-12-22
:	:	: PRIOR APPLICATION NUMBER: 60/269,308
:	:	: PRIOR FILING DATE: 2001-02-16
:	:	: NUMBER OF SEQ ID NOS: 14110
:	:	: SOFTWARE: FastSEQ for Windows Version 4.0
:	:	: SEQ ID NO I3083
:	:	: LENGTH: 1048
:	:	: TYPE: PRT
:	:	: ORGANISM: Staphylococcus aureus
:	:	: US-09-815-242-13083
Query Match 8.1%; Score 162; DB 9; Length 1048;		
Best Local Similarity 23.8%; Pred. No. 0.00035;		
Matches 88; Conservative 53; Mismatches 146; Indels 82; Gaps 19;		
QY	17	TLEAAARWELTDLINAAMTILASIQDYAKTEASLSAYSEA-EVTNNLNATLQLKMAK 75
Dd	85	TDLNTAGNLOGAINDEQTLLNS-QNYQDATPSKKTYATNAVQAADILNKSGNQK-TK 142
QY	76	TNLESAINQAWTKTTEDNEHPNLVEAYKALKTTLEQRATNLEGSLSTAYNQIRNNLYDL 135
Dd	143	DQVTEAMNVNSAKNNLDG-----TRLIDQAKOTAKOOLNNMTHLTTAOKTNLTNQ---- 193
QY	136	YNKASSLITKTLDPLNGTGTLDSNEITTANKN-INNTLSTINEQKTNADA--LSNSFI-- 190
Dd	194	-----INSGETTVAGVQTVGSNANTLDAQMTLRQSIAKDASKASYDVA 239
QY	191	---KKVIQNNEPQSFVGITFNANOPFS-NYSVFASFADVTPPNKYARTVNWGE----- 241
Dd	240	NNDQGTAYNNAVAATAETIINASNPEMPNPSTITOKAE--QVN---SSKALTGNENLAFA 294
QY	242	-PPSRILANT-NSITD-----VSWIYS---LAGNTKYQFSPSNYGPGSTGYLPYPKLIV 290
Dd	295	KONAKTYLNTLTSITDAQKNLLISOITSATRVSVDVTVKQNA----- 336
QY	291	KAADANNNGLOVCKLNGNVQQVEAFATS-SANNTTTANPTPAVDE-IKVAKIVL---SGLRF 346
Dd	337	QHLDQMASLQNGINN-----ESQVKSSEKRYRADNTNKOQEYDNATAAKAILNKGSTGPNT 392

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:28:16 ; Search time 29 seconds
(without alignments)
574.844 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTKKDANPNNGTQLE.....SSNENNADKIPGRRPGTFL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	95.3	615	2	US-08-525-742-10
2	1856	92.6	610	2	US-08-525-742-8
3	1638	81.7	368	1	US-08-525-742-4
4	1612	80.4	368	1	US-08-185-851A-4
5	806	40.2	661	2	US-08-525-742-2
6	175	8.7	10182	4	US-09-134-001C-3159
7	159	7.9	933	3	US-08-293-728-2
8	159	7.9	933	3	US-09-421-868-2
9	154.5	7.7	682	4	US-08-836-687B-37
10	136	6.8	1098	3	US-08-923-992A-8
11	135.5	6.8	1002	4	US-09-268-347-24
12	135.5	6.8	1073	3	US-09-541-782-6
13	135.5	6.8	1073	4	US-09-723-820-6
14	134	6.7	1004	4	US-09-268-347-30
15	131.5	6.6	1184	3	US-08-923-992A-2
16	131	6.5	1104	3	US-08-923-992A-4
17	129	6.4	1833	4	US-08-621-944A-4
18	129	6.4	1833	4	US-08-945-567D-4
19	129	6.4	1992	4	US-08-621-944A-3
20	129	6.4	1992	4	US-08-945-567D-3
21	129	6.4	2048	4	US-09-268-347-48
22	128	6.4	1095	4	US-09-206-942-45
23	128	6.4	1101	4	US-09-206-942-43
24	127.5	6.4	2504	4	US-09-328-352-5821
25	126.5	6.3	1128	3	US-08-923-992A-6
26	126.5	6.3	1164	3	US-08-923-992A-10
27	125.5	6.3	518	3	US-09-043-123-2

28	125.5	6.3	1073	4	US-09-206-942-49	Sequence 49, Appl
29	125.5	6.3	1079	4	US-09-206-942-47	Sequence 47, Appl
30	125.5	6.3	3788	4	US-09-336-447A-76	Sequence 76, Appl
31	124.5	6.2	2411	4	US-09-268-347-36	Sequence 36, Appl
32	123	6.1	969	4	US-09-206-942-32	Sequence 32, Appl
33	123	6.1	975	4	US-09-206-942-30	Sequence 30, Appl
34	122.5	6.1	1095	4	US-09-206-942-69	Sequence 69, Appl
35	122.5	6.1	1536	1	US-08-038-682-2	Sequence 2, Appl
36	122.5	6.1	1536	1	US-08-302-832-2	Sequence 2, Appl
37	122.5	6.1	1536	2	US-08-530-198-2	Sequence 2, Appl
38	122.5	6.1	1536	2	US-08-469-880-2	Sequence 2, Appl
39	122.5	6.1	1536	2	US-08-728-470-2	Sequence 2, Appl
40	122.5	6.1	1536	2	US-08-617-697-2	Sequence 2, Appl
41	122.5	6.1	1536	3	US-08-719-641-2	Sequence 2, Appl
42	122.5	6.1	1536	4	US-09-206-942-67	Sequence 67, Appl
43	122.5	6.1	1565	6	5352450-2	Patent No. 5352450
44	122	6.1	690	4	US-09-134-001C-4568	Sequence 4568, Ap
45	122	6.1	2137	4	US-09-134-001C-4463	Sequence 4463, Ap

ALIGNMENTS

RESULT 1
US-08-525-742-10
; Sequence 10, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Sakai, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroh
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA: PCT/JP94/00541
; APPLICATION NUMBER: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930

```
TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-10

Query Match          95.3%; Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%; Pred. No. 1.4e-131;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTOLKARMLTDLINAKAMTLASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDKANPNNGQTOLKARMLTDLINAKAMTLASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLA 146
QY 122 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQKTN 181
DB 147 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQKTN 206
QY 182 ADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRVWNGDE 241
DB 207 ADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRVWNGDE 266
QY 242 PSSRLANTNSITDVSWIYSLAGNTKYQFSNYPSTGYLYPPYKLVKAADANNVGLQ 301
DB 267 PSSRLANTNSITDVSWIYSLAGNTKYQFSNYPSTGYLYPPYKLVKAADANNVGLQ 326
QY 302 YKLNGNVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 361
DB 327 YKLNGNVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSNNENADKI 384
DB 387 MNKVAPMIGNIYLSNNENADKI 409

RESULT 2
US-08-525-742-8
; Sequence 8, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroh
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyochito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA: PCT/JP94/00541
; APPLICATION NUMBER: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-859-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-8

Query Match          92.6%; Score 1856; DB 2; Length 610;
Best Local Similarity 95.5%; Pred. No. 1.2e-127;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTOLKARMLTDLINAKAMTLASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDKANPNNGQTOLKARMLTDLINAKAMTLASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 146
QY 122 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQKTN 181
DB 147 STAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQKTN 206
QY 182 ADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRVWNGDE 241
DB 207 ADALNSFIKKVIONNEOSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRVWNGDE 266
QY 242 PSSRLANTNSITDVSWIYSLAGNTKYQFSNYPSTGYLYPPYKLVKAADANNVGLQ 301
DB 267 PSSRLANTNSITDVSWIYSLAGNTKYQFSNYPSTGYLYPPYKLVKAADANNVGLQ 326
QY 302 YKLNGNVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 361
DB 327 YKLNGNVQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSNNENADKI 383
DB 387 MNKVAPMIGNIYLSNNENADKI 408

RESULT 3
US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroh
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
```


QY 302 YKLNNGVQOEFATSTANNTTANPTPAVDEIKVAK 338
 Db 327 YKLNNGVQOEFATSTANNTTANPTQOOLMLKLIK 363

RESULT 5
 US-08-525-742-2
 ; Sequence 2, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hiroko
 ; APPLICANT: Iritani, Yoshihiko
 ; APPLICANT: Aoyama, Shigemi
 ; APPLICANT: Takahashi, Kiyohito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSEE: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742
 ; FILING DATE: 25-SEP-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-074139
 ; FILING DATE: 31-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-245625
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP94/00541
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLeland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 950811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-659-2930
 ; TELEFAX: 202-8870357
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-525-742-2

Query Match 40.2%; Score 806; DB 2; Length 661;
 Best Local Similarity 42.8%; Pred. No. 6e-51;
 Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;
 QY 9 DANPNNGQQLAEARMELTD-----LINAKMTLASIQDYAKTEASLSAYSEATVN 61
 Db 45 NTNPSDGGMMNAKADAKAALTTLINGETANLASEYAKIKSELTSAKETAKAYS 104
 QY 62 NNUNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVYAKALKTTLEQRTNLEGLS 121

Db 105 AKTGATLNEVNEAKTTLDAAIKKAASAKNDFDAQHGLSLVEAYNNKLTKEETNLSLA 164
 QY 122 STAYNOIRNNLVLYNKASSLITKTLDPLNGGTLDSNEITTANKNNINNTLSTINEOKTN 181
 Db 165 NENTAAITNLSLYEKANTIVTATLDPAT-GNIPVMSVTOANQDITNATSLRIANKQN 223
 QY 182 ADALNSFIKVIQIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVPVYKARRTVWNGDE 241
 Db 224 ADNLANSFIKQSLVKNNLTRVDVANNOE-OPANTSFVGFVSYNVDTPNWNFAQRKVASEN 282
 QY 242 -----PSSRLANTNSITDYSWLYSLAGTNTKYQFSFNSYG-PSTGVLYFPYKLVKAAAD 294
 Db 283 TPLATTPAEDATQQAASLTDVSWIYSLNGABAKYTLSTFRFGAETAYLYFPYKLVKTS 342
 QY 295 ANNVGLOYKLNNGVQOEF-----ATTSAN--NTTANP-----TPAVDEIKVAKIV 340
 Db 343 --NVGLOYKLNNGGDKQINFVQTPASGSSDVAANEETMASPAEQSAPTVDIKIAKVA 400
 QY 341 LSLGRFGQNTIELSVPTGEGNNKVPAMIGNIYLS-----NNNN 380
 Db 401 LSLNLFNSNTIEFSYPTG-----KAAPMIGNMYLTSSNSEVNNKNIYDLDLFGSNFNNEN 455
 QY 381 -----ADKIPGY 387
 Db 456 PTAVTVDLLKGY 467

RESULT 6
 US-09-134-001C-3159
 ; Sequence 3159, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3159
 ; LENGTH: 10182
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-3159

Query Match 8.7%; Score 175; DB 4; Length 10182;
 Best Local Similarity 20.1%; Pred. No. 0.003;
 Matches 114; Conservative 71; Mismatches 182; Indels 200; Gaps 22;
 QY 8 KDANPNNGQT-----QLEARMELTD-----LINAK-----AM 35
 Db 3980 ENSNPSVNEVAQALQKVEAVQLKVNDAIHLQNKENSALVTAKNLQOQSVNDQPLTTGM 4039
 QY 36 TLASLDYAKIEASLSAYSEATVNNLNATLEQLKMAK-----75
 Db 4040 TQDSINNYEAKRNEAQSATRNAEAVINNGDATAKQISDEKSKVEQALAHNDKQOITAD 4099
 QY 76 -TNLESAINQANTDKTTFDNEHPNLVYAKALKTTLEQRTNLEGLSSTAYNQIR-----129
 Db 4100 TTELQTAVALQNLNRRGDT--NNKKPRSIYAKAQSLEQTTSKADNANAVIQRPITVQE 4158
 QY 130 -NNVLDVLYNKASSLIYKT---LDPLNGGTLDSNEITTANKNNINNTLSTINEOKTNADAL 185
 Db 4159 VNNALQOVNLNOQLTEAINOLOPLSNNDALKARLNLENK-INQTVQTDGCTQOQSLEY 4217
 QY 186 SNSFIKVIQIONNEQSFVGTFTN-----ANVQPSNY 215
 Db 4218 QNA--KRAQNESNTALALNNGDADQOITETDTRYNOQTNTLQTOALINGLTVNKEPLET 4275

QY 216 SFVAFSADVTPV-----NKVARRTVWNGDEPSSRLAN----- 249
DB 4276 AKTALQNNIDQVPSTGDTGTOQSVANYNKLQAKNEINTINNVLANNPDVNAIKTNKABE 4335
QY 250 -----TNSIT-----DVSWIYSL-----AGTNT--KYQFSFSNYGPS--TGYLYFP 286
DB 4336 ERISNDLTOAKNNLQVDTOPLEKIKROLQDEIDQGTNDGTQDSVDNYNDSLSAAIEK 4395
QY 287 YKLVAADANNVGLQKLNNGVQV----- 312
DB 4396 GKNVLLKRNPTVEQVKESVANAQVQIDLOQARTSLVPDKTQLQEAKNRLNSINQOQT 4455
QY 313 -FEATSTANNANTPAVDI-KVAKIVLSGLRFGONTI-ELSVPTGEGNNKVA--- 366
DB 4456 TCGMTQDSLNNYNDKLAKRQNLKISKVL-----GGQTVAEIRQNTDEANAHQALDT 4510
QY 367 -----PMIGNIYLSNNENNAK 383
DB 4511 ARSOLTNRPEVYNINHNESHLNNAK 4537
RESULT 7
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 7.9%; Score 159; DB 3; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.0016;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;
QY 3 MSITKDPNNGQOTL--EARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATV 60
DB 73 VSDTKTSSNTNGETSVAGNPAQOQETQSSSTNAIT-----EETPVTGEATTT 120
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAINLEGL 120
DB 121 TTNOANTPATTOSSNTNAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
QY 121 SST-----AYNQIRNNLVLYNK-----ASSLITKTLPLNGGTL 156
DB 173 STQDTSTATPSNNEAPQSDASKNOVVNOAVNTSAPRMRAFSLAAVAADAPAGTDI 232
QY 157 DSNEITANKNNITLSTINEQ-----KTN-ADALNSFIK-----KVIONNEQSFVGTFTN 207
DB 233 -TNQLNTVIGDSGTVVPHQAGYVKLYGFSVPNSAVKGTFTKIVPKELNNGVTST 291
QY 208 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDYSWIYSLAG--TN 266
DB 292 AKVPP-----IMAGDQ-----VLANGVIDSDGNVIYTFDYVN 324
QY 267 TTYQFSFSGYNGPSTGYLYPPYKLVKRAADANNVGLQKLNNGVQVFEFATSTANNNTAN 326
DB 325 TRDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
QY 327 PTPAVDEIKVAKI-----VLSGLRFGONTIELSVPTGEGNNKVPAMI-GNIYLSNE 378
DB 363 KTVLVDYKYGKFNLSIKGTIDQDKTNTYRTQTIYVNPSPDGNVIAPVLTGKLNKPTDS 422

QY 379 N 379
DB 423 N 423
RESULT 8
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 7.9%; Score 159; DB 3; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.0016;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;
QY 3 MSITKDPNNGQOTL--EARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATV 60
DB 73 VSDTKTSSNTNGETSVAGNPAQOQETQSSSTNAIT-----EETPVTGEATTT 120
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAINLEGL 120
DB 121 TTNOANTPATTOSSNTNAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
QY 121 SST-----AYNQIRNNLVLYNK-----ASSLITKTLPLNGGTL 156
DB 173 STQDTSTATPSNNEAPQSDASKNOVVNOAVNTSAPRMRAFSLAAVAADAPAGTDI 232
QY 157 DSNEITANKNNITLSTINEQ-----KTN-ADALNSFIK-----KVIONNEQSFVGTFTN 207
DB 233 -TNQLNTVIGDSGTVVPHQAGYVKLYGFSVPNSAVKGTFTKIVPKELNNGVTST 291
QY 208 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDYSWIYSLAG--TN 266
DB 292 AKVPP-----IMAGDQ-----VLANGVIDSDGNVIYTFDYVN 324
QY 267 TTYQFSFSGYNGPSTGYLYPPYKLVKRAADANNVGLQKLNNGVQVFEFATSTANNNTAN 326
DB 325 TRDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
QY 327 PTPAVDEIKVAKI-----VLSGLRFGONTIELSVPTGEGNNKVPAMI-GNIYLSNE 378
DB 363 KTVLVDYKYGKFNLSIKGTIDQDKTNTYRTQTIYVNPSPDGNVIAPVLTGKLNKPTDS 422
QY 379 N 379
DB 423 N 423
RESULT 9
US-08-836-687B-37
; Sequence 37, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B

```

; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 682
; TYPE: PRN
; ORGANISM: Lactococcus sp.
US-08-836-687B-37

```

Query Match	7.7%	Score 154.5;	DB 4;	Length 682;
Best Local Similarity	21.0%;	Pred. No. 0.0023;		
Matches	75;	Conservative 61;	Mismatches 130;	Indels 91; Gaps 13;

Qy	58	ETVNNLNATLEQLKMAKTNLGSAINQANTDKTTFEDNEHPNLVAYKALKTTLEQRATNL	117
Db	38	ELIHNHNSALLSTEGSTTYSINLGAQSPAVKSYTRTE----	177
Qy	118	EGLSSTAYNQIRNNLDVLYNKRASSLIKTLDPLNGGTLDSNETTTANKNINNTLSTINE	177
Db	94	EMKVSLEQTQVSSE---FSKRDSVTNKEAVPVSKDELLEQSEVVVSTSSIOKN-KILDN	148
Qy	178	QKTNADALSNFSFKKVLQNNHQEQSFVGTFTNANQVPSYKSVAFPSADTVPYNYKKTARRTV-	236
Db	149	KKRANFVTSPLTIKEXPSNKSQASGVIDN-----ASPLSYRKAKEVYS	193
Qy	237	-----WNGDEPSSRILANTINSITDVSWIYSLAGTNTKYQFSFSNYGPS	279
Db	194	LROPLKNQKVEAQPPLLINSSEKKASYTTNSHDFWDYQW-----DMKY---VTNNGES	243
Qy	280	TGVLFFPKVLKKAADANNVGLQYKLNNGNQVQEFATSTSAN-----NT	323
Db	244	YA-LYQPSKKI-----SVGI-----IDSGIMEEHFDLSNGLYFNKLVPKGFGDNEEDE	293
Qy	324	TANPTPAVDIEIKVAKIVLSGLRFQONTIEIUSVPTGEGNMNKVAPMIG-NIYLSNNEN	379
Db	294	TGNPSDITD-----KMGHGTVEAGQITANGNIGLVAQGITNIIYRVFGEN	338

RESULT 10
 US-08-923-992A-8
 ; Sequence 8, Application US/08923992A
 ; Patent No. 6280738
 ; GENERAL INFORMATION:
 ; APPLICANT: Tai, Joseph Y.
 ; APPLICANT: Blake, Milan S.
 ; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B B
 ; TITLE OF INVENTION: Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 34
 ; ADDRESS/SEQUENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/923,992A
 ; FILING DATE: 05-SEP-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/024,707
 ; FILING DATE: 06-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600

```

; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

```

Query Match	6.88;	Score 136;	DB 3;	Length 1098;
Best local Similarity	22.7%;	Pred. No. 0.098;		
Matches	84;	Conservative 51;	Mismatches 159;	Indels 76; Gaps 14;
Qy	5	ITKRDAMPNGQTGLEAARMELTDLINA-KAMTLIASLDQYAKIEASLSAYSE-----AE	58	
Db	192	IRKQAOQPKKEDAEVKVREELKLFSTRAGLDQEIQEHVKKETSSEENTQKVDEHYAN	251	
Qy	59	TVNNLNATYLEOLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALK	107	
Db	252	SLONLAKOSLEELD KATN-----EQATQVKNQFLENAQKLKEIQLPIKETNVKLYKAMS	306	
Qy	108	TTLEORATNLEGLSSTAYNQIRNNLVLYNKASSLIITKLDPLNGGTLTLDNSNETTTANKN	167	
Db	307	ESLEQVEKELEKHNSEA-----NLEDLVAKSEIVREYEGKLNQSKNLP--ELKQLEEE	357	
Qy	168	INNTLSTINEQ-----KTNADALNSNFIKKVI-----QNNQESFVGTFTNANVPQSNYSFVA	219	
Db	358	AHSLKQVWEDFRKKFETSEQVTPKKRLKRDLAANENNQKI-----ELTVSPENI-----	408	
Qy	220	FSADVTVPVNYKARRTTWNGDEPSSRIILANTINSITDYSWISLAGTWTQYQFSF-----	274	
Db	409	-----TYVEGEDVKFTVTAKSQSKTTLDFSDLL-----TKYNPWSDRIST	449	
Qy	275	NYGPSTG---YLPPYKLVKAADANNVGLQYKLNNGNVQVQFEFATSTSANNTTANP--TPA	330	
Db	450	NYKTNTDNHKAETAITIKNLKINESQVTYLKAKDDSGNVVEKFTFTITVQKKEQVKPTKE	509	
Qy	331	VDEIKVAKIV	340	
Db	510	QKDSKTEEKV	519	

```

RESULT 11
US-09-268-347-24
; Sequence 24, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

```

	Query Match	6.8%;	Score 135.5;	DB 4;	Length 1002;
	Best Local Similarity	19.2%;	Pred. NO. 0.094;		
	Matches	94;	Conservative 60;	Mismatches 161;	Indels 175; Gaps 21
QY	1	GCMSTTKDANPNNGOTOLEARM----	ELDTLINAKAMTLASLODYAKIEASLSAYSE	56	
		:: :: :: :: :: :: :: :: ::			
Db	349	GKLARTKVKLVSGANGTNPVKSINVDGTEDDAVSFK--QLKALQD-KOVLTSASNAYAN	405		
QY	57	-----AETVNNNLNATLEOLKMAKTNLESA-	-INQANTDKTTEDN	94	
		:: :: :: :: :: :: :: :: ::			
Db	406	GGSDADGGKATQTGLNDLNFKEFSTDSELLNIKAAGDVTFTPPKGSVQVGDDGRATQD	465		
QY	95	-----EHPNLVEAYKAL-----K TTL-----	110		

Db 466 GAKTTTGLVEASELVDLSNKLGWKGVGKDGATGATDGTHTDTLVKSGDKVTLKAGDNLV 525
Qy 111 EQRATN-----LEGSSSTAYNQIRNNLVLYNKASSLTK-----TLDP---LNGGTLL 156
Db 526 KQECTNFTYVLRDELTCVKSVFKDTENGA---NGASTKITKDLGITTPANDANGAAAT 581
Qy 157 DSNBITTANKNINNTLSTINEQKTNADALSNSFTKKYQVIONEQSFVGTFTNANVOPSNYS 216
Db 582 DADRIKVASDGI-----SAGNKAVKNVY-----SGLKKFGDANFN----- 617
Qy 217 FVAFSADVTVPVNYKARTVWNGDEPS-----SRILANTNSITDVSIIYS---LAG 264
Db 618 -LTSADNLTKQYDNAYKGLTNLDEKSGKQTPPVADNTAATVGDRLGLGWISADTKTG 676
Qy 265 TNRKYQFSFGSTGYLFPFKVLAADANNVGLQYKLNNGNVQVQVEFATISANNTT 324
Db 677 ESKEYSAQVRNANE-----VKFKSGINGVSGKTLNDGTREITPELAKDEN--- 722
Qy 325 ANPTPAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVAPMI 369
Db 723 -----AIAFGSGKALRDNTVAIGTGVNVNAEKSGAFGDPNYIEDKA 764
Qy 370 GNIYSSNEN 379
Db 765 GGSTAFGNDN 774

RESULT 12

US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541.782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6

Query Match 6.8%; Score 135.5; DB 3; Length 1073;
Best Local Similarity 20.9%; Pred. No. 0.1;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;
Qy 14 NGQ-TQLEAARMELTDLINAKA-MTLASL-----ODYAKIEASLSAYSEAEVTNNLNAT 67
Db 611 NGYFTLLDNFNASMEELLNTHSNOLLISMTKITEHFQSLDEALQASRSCAVPNSSLDLI 670
Qy 68 LEQLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEQRATNLEG 119
Db 671 VSELKDSKNSLLDALEHSLQDISMSQKLGNGISSELIELQDKMESYRQLVQELRSLYN 730
Qy 120 LSSSTAYNQIRNNLVLYNKASSLITKTLDPNGTTLDSNEIT-----TA 164
Db 731 LQHTHEESQKELMTGVRNDIDALVKTCTTSLNDADIILSDYISDQKSFESKQDOLIANI 790
Qy 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 214
Db 791 GKIVSNFLQEQNESLYTKADILHSLNDNTSNIRKANEIMNRSEFL---RNA----- 841
Qy 215 YSFVAFSADVTVPVNYKARTVWNG-----DEPSSRILANTNSITDVSIIYSLA----- 263
Db 842 -----ASQAEIVGANKERIQKTVEGSQLDLSKSAIHNSRSMYD-----HCLALAESQK 893
Qy 264 GTNTKYQ-----FSFSNYGPSTGYLFPFKVLAADANNVGL---OYKLNNGNVQ 310
Db 731 LQHTHEESQKELMTGVRNDIDALVKTCTTSLNDADIILSDYISDQKSFESKQDOLIANI 790
Qy 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 214
Db 791 GKIVSNFLQEQNESLYTKADILHSLNDNTSNIRKANEIMNRSEFL---RNA----- 841
Qy 215 YSFVAFSADVTVPVNYKARTVWNG-----DEPSSRILANTNSITDVSIIYSLA----- 263
Db 842 -----ASQAEIVGANKERIQKTVEGSQLDLSKSAIHNSRSMYD-----HCLALAESQK 893
Qy 264 GTNTKYQ-----FSFSNYGPSTGYLFPFKVLAADANNVGL---OYKLNNGNVQ 310

Db 894 GVNLEVTQLRLKQVKEHSEDNTKEKHQQL---LDLLESVGNNDNLIDSIKTPHTELQ 950
Qy 311 QVE-----FATTSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 351
Db 951 KITDHLVKGTTSLANHTNELLGLGDESICNLNETTIEDTSLVKLETTGDTTPSKRELPAPTS 1010
Qy 352 -----ELSVPTGEGNMKNVAPMIGNIYSSNENNAKIPGYRRP 390
Db 1011 WTRDSSLIKETTNNLSDSKKFRVRETYTSSNQTNPEDV--YDKP 1052
RESULT 13
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723.820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541.782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-723-820-6

Query Match 6.8%; Score 135.5; DB 4; Length 1073;
Best Local Similarity 20.9%; Pred. No. 0.1; Mismatches 182; Indels 109; Gaps 21;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;
Qy 14 NGQ-TQLEAARMELTDLINAKA-MTLASL-----ODYAKIEASLSAYSEAEVTNNLNAT 67
Db 611 NGYFTLLDNFNASMEELLNTHSNOLLISMTKITEHFQSLDEALQASRSCAVPNSSLDLI 670
Qy 68 LEQLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEQRATNLEG 119
Db 671 VSELKDSKNSLLDALEHSLQDISMSQKLGNGISSELIELQDKMESYRQLVQELRSLYN 730
Qy 120 LSSSTAYNQIRNNLVLYNKASSLITKTLDPNGTTLDSNEIT-----TA 164
Db 731 LQHTHEESQKELMTGVRNDIDALVKTCTTSLNDADIILSDYISDQKSFESKQDOLIANI 790
Qy 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 214
Db 791 GKIVSNFLQEQNESLYTKADILHSLNDNTSNIRKANEIMNRSEFL---RNA----- 841
Qy 215 YSFVAFSADVTVPVNYKARTVWNG-----DEPSSRILANTNSITDVSIIYSLA----- 263
Db 842 -----ASQAEIVGANKERIQKTVEGSQLDLSKSAIHNSRSMYD-----HCLALAESQK 893
Qy 264 GTNTKYQ-----FSFSNYGPSTGYLFPFKVLAADANNVGL---OYKLNNGNVQ 310
Db 894 GVNLEVTQLRLKQVKEHSEDNTKEKHQQL---LDLLESVGNNDNLIDSIKTPHTELQ 950
Qy 311 QVE-----FATTSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 351
Db 951 KITDHLVKGTTSLANHTNELLGLGDESICNLNETTIEDTSLVKLETTGDTTPSKRELPAPTS 1010
Qy 352 -----ELSVPTGEGNMKNVAPMIGNIYSSNENNAKIPGYRRP 390
Db 1011 WTRDSSLIKETTNNLSDSKKFRVRETYTSSNQTNPEDV--YDKP 1052
RESULT 14
US-09-268-347-30

; Sequence 30, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1004
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-268-347-30

Query Match 6.7%; Score 134; DB 4; Length 1004;
Best Local Similarity 20.3%; Pred. No. 0.12;
Matches 100; Conservative 53; Mismatches 160; Indels 180; Gaps 22;

QY 1 GCSITKKDANPNNGQTQLEARM-----ELTDLINAKAMTSLASLDYAKIEASLSAYSE 56
Db 346 GKLAQKVLVLSANGPNVKISNVADGTDNDVSEK--OLKALQD-KQVTLASNAYAN 402
QY 57 A-----ETVNNLNATLEQLKMAKTINLESAINQANTDKTF-----DN-- 94
Db 403 GGSADAGGKGIQTLNGLN-----FKFKSTDGELLNKAENDIVTTPKGSVQVGDGK 457
QY 95 -----EHPNVLVEAYKAL-----KTTL--- 110
Db 458 ATIQDGAQKTTGLVEASELVDLSLNLKLGKVGVTGDTGVTGTHDTLVKSGDKVTLKAG 517
QY 111 -----BORATN-----LEGSSPAYNQIRNLDVLYNKASSLTK---TLDPLNGGT 154
Db 518 DNLKVKQEGNTFYALKDELTDVKSVEFKDTANGA-----NGASTKITKGLTITPANGAG 573
QY 155 LLDSEITANKINNLTSTINEQRTNADALSNSFIKKVIQNNQSEFVGTFTNANVQPSN 214
Db 574 AAGA-----NTANTISVTKDGIS-A-GNKAVKNV-----SGLKKFGDANFDP-- 614
QY 215 YSFVPSADVTPVNYKYARTVWNGDEPS-----SRILANTNSITDVSWIYSLAGT 265
Db 615 ----LTSSADNLTKQYDNAYKGLTNLDEKSGKQTPVADNTAATVGDGLRGLGVISADKT 671
QY 266 ----NTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVQYKLNNGNVQVQVEFATSTAN 321
Db 672 KGLNKEYNAQVRANE-----VFKSGNGINVGSKTLDNGTRETITPELAKDEN 720
QY 322 NTTANPTPAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVA 366
Db 721 -----AIAFGSGSKALRDNTVAIGTGNVNVNAEKSGAFGDPNYIE 759

QY 367 PMIGNIYLSNEN 379
Db 760 DRAGGSYAEGNDN 772

RESULT 15
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-923-992A-2

Query Match 6.6%; Score 131.5; DB 3; Length 1164;
Best Local Similarity 22.8%; Pred. No. 0.23;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY * 7 KKDANPNNGQTQLEA-ARMELTDLINA-KAMTSLASLDYAKIEASLSAYSE-----AET 59
Db 230 RQAOQAQADKEDAEVKEELGKLFSTKAGLDQEIQEHVKETSSSEENTQKYDEHYANS 289
QY 60 VNNLNATLEQLKMAKTINLESAINQANTDKTTFDNEHPLN-----VEAYKALKT 108
Db 290 LQNLAKSLEELDKATTN-----EQATQVKNQFLENAQKLKEIPLIKETNVKLYKAMSE 344
QY 109 TLEQRTNLEGSSPAYNQIRNLDVLYNKASSLITKTLDPNGGTLSDSNEITTANKNI 168
Db 345 SLEQVEKELKHNSA-----NLEDLVAKSKEIVREYEGKLNQSKNLP--ELKQLEEEA 395
QY 169 NNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQSFVGTFTNANVQPSNYSFVAF 220
Db 396 HSKLKQVVEDFRKKFTSQVTPKRVKRDLAANNENQOKI-----ELTVSPENI----- 445
QY 221 SADVTPVNYKYARTVWNGDEPSRILANTNSITDVSWIYSLAGTNTKYQFSFS-----N 275
Db 446 -----TYEGEDVKFTVAKSDSKTTLDFSDLL-----TKYNPSVSDRISTN 487
QY 276 YGPSTG---YLFPYKLVKAADANNVQYKLNNGNVQVQVEFATSTANNTTANP-TPAV 331
Db 488 YKINTDNHKAIEITIKNLKLNESQVTLRAKDDSGNVVKEFTTITVQKKEEKQVKTPEQ 547
QY 332 DEIKVAKIV 340
Db 548 KDSKTEEV 556

Search completed: October 8, 2003, 17:33:12
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:08:36 ; Search time 82 Seconds

(without alignments)
762.662 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTTKDANPNNGQTGLE.....SSNNENADKIPGRRPGTFL 394

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2004	100.0	456	18 AAW36050	Hybrid Marek's dis
2	2000	99.8	1086	18 AAW36051	Hybrid Marek's dis
3	1910	95.3	615	15 AAR63230	Mycoplasma gallise
4	1856	92.6	610	15 AAR63229	Mycoplasma gallise
5	1612	80.4	368	14 AAR44493	Mycoplasma gallise
6	1612	80.4	368	15 AAR63227	Mycoplasma gallise
7	1604	80.0	368	16 AAR76955	Mycoplasma gallise
8	1142	57.0	235	10 AAP93646	Amino acid sequenc
9	1142	57.0	235	11 AAR05081	MG-1 antigen. Myc

10	1142	57.0	261	11 AAR05082	TMG-1 antigen. My
11	1142	57.0	261	16 AAR79911	M.gallisepticum 26
12	1117	55.7	261	10 AAP93959	Amino acid (AA) se
13	806	40.2	661	15 AAR63226	Mycoplasma gallise
14	806	40.2	661	16 AAR79910	M.gallisepticum 66
15	713	35.6	647	16 AAW11978	Mycobacterium gall
16	700.5	35.0	648	15 AAW56973	PMGA 1.2 protein o
17	386.5	19.3	183	10 AAP93649	Amino acid sequenc
18	386.5	19.3	183	11 AAR06439	MG-4 antigen. Myc
19	284.5	14.2	219	10 AAP93648	Amino acid sequenc
20	284.5	14.2	219	11 AAR06438	MG-3 antigen. Myc
21	176.5	8.8	6281	22 AAO37403	Staphylococcus aur
22	176	8.8	1095	22 AAG83030	S. epidermidis ope
23	175	8.7	10182	23 ABP38314	Staphylococcus epi
24	174.5	8.7	10498	24 ABJ19119	Pathogen specific
25	171	8.5	2086	22 AAO34143	Staphylococcus aur
26	171	8.5	5795	22 AAO37017	Staphylococcus aur
27	162.5	8.1	2659	23 AAO75489	S. aureus antigeni
28	162	8.1	1029	22 AAO34389	Staphylococcus aur
29	162	8.1	1048	22 AAO37490	Staphylococcus aur
30	160	8.0	933	24 ABJ18947	Pathogen specific
31	159	7.9	496	23 AAO75490	S. aureus antigeni
32	159	7.9	520	23 AAE29262	Staphylococcus aur
33	159	7.9	933	21 AAY58435	Staphylococcus aur
34	159	7.9	933	22 AAB69508	Staphylococcus aur
35	159	7.9	936	18 AAW89801	Staphylococcus aur
36	158	7.9	2434	22 AAO34339	Staphylococcus aur
37	156	7.8	807	21 AAB18311	Plasmodium falcipa
38	155	7.7	1072	23 ABB54963	Lactococcus lactis
39	154.5	7.7	682	17 AAR95273	Nisin nisp gene pr
40	153.5	7.7	2478	22 AAO34320	Staphylococcus aur
41	153.5	7.7	2478	22 AAO37374	Staphylococcus aur
42	153.5	7.7	2478	24 ABJ19002	Pathogen specific
43	151.5	7.6	1237	18 AAW55640	H. pylori ORF 04ep
44	151.5	7.6	1237	20 AAY17187	H. pylori outer me
45	151	7.5	5024	22 AAG82935	S. epidermidis ope

ALIGNMENTS

RESULT 1

ID AAW36050 standard; Protein; 456 AA.

XX AAW36050;

XX 15-JUL-1998 (first entry)

XX Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein; antigen; vaccine; poultry.

XX Chimeric - Marek's disease gammaherpesvirus.

OS Chimeric - Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Region 1..64

FT Region /note= "derived from Marek's disease virus gb protein"

FT Region 65..456

FT Region /note= "derived from M. gallisepticum antigenic protein"

XX WO9736924-A1.

PN 09-OCT-1997.

XX 28-MAR-1997; 97WO-JP01084.

XX 29-MAR-1996; 96JP-0103548.

XX (JAPG) NIPPON ZEON KK.

PI Saito S, Tsuzaki Y, Yanagida N;
 XX WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 456 AA;
 Query Match 100.0%; Score 2004; DB 18; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.8e-128; Indels 0; Gaps 0;
 Matches 394; Conservative 0; Mismatches 0;
 QY 1 GCMSTTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 60
 DB 63 GCMSTTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 122
 QY 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
 DB 123 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 182
 QY 121 SSTATNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTKANKNINNTLSTINEQKT 180
 DB 183 SSTATNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTKANKNINNTLSTINEQKT 242
 QY 181 NADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGD 240
 DB 243 NADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGD 302
 QY 241 EPSSRLANTNSITDVSIIYSLAGNTKYQSFSGNYPSTGYLYPPYKLVKAADANNVGL 300
 DB 303 EPSSRLANTNSITDVSIIYSLAGNTKYQSFSGNYPSTGYLYPPYKLVKAADANNVGL 362
 QY 301 QYKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
 DB 363 QYKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
 QY 361 NNNKVAPMIGNIYLSNENNADKIPGYRRPGTFL 394
 DB 423 NNNKVAPMIGNIYLSNENNADKIPGYRRPGTFL 456
 RESULT 2
 AAW36051
 ID AAW36051 standard; Protein; 1086 AA.
 XX
 AC AAW36051;
 XX
 DT 15-JUL-1998 (first entry)
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 XX
 KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 XX
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Region 1..672
 FT /note- "derived from Marek's disease virus gB protein"
 FT 693..1086

FT
 XX
 PN W09736924-A1.
 XX
 PD 09-OCT-1997.
 XX
 PF 28-MAR-1997; 97WO-JP01084.
 XX
 PR 29-MAR-1996; 96JP-0103548.
 XX
 PA (JAPG) NIPPON ZEON KK.
 XX
 PI Saito S, Tsuzaki Y, Yanagida N;
 XX
 DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96596.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;
 Query Match 99.8%; Score 2000; DB 18; Length 1086;
 Best Local Similarity 99.7%; Pred. No. 1.1e-127;
 Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCMSTTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 60
 DB 693 GCMSTTKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 752
 QY 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
 DB 753 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 812
 QY 121 SSTATNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTKANKNINNTLSTINEQKT 180
 DB 813 SSTATNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTKANKNINNTLSTINEQKT 872
 QY 181 NADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGD 240
 DB 873 NADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGD 932
 QY 241 EPSSRLANTNSITDVSIIYSLAGNTKYQSFSGNYPSTGYLYPPYKLVKAADANNVGL 300
 DB 933 EPSSRLANTNSITDVSIIYSLAGNTKYQSFSGNYPSTGYLYPPYKLVKAADANNVGL 992
 QY 301 QYKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 360
 DB 993 QYKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 1052
 QY 361 NNNKVAPMIGNIYLSNENNADKIPGYRRPGTFL 394
 DB 1053 NNNKVAPMIGNIYLSNENNADKIPGYRRPGTFL 1086
 RESULT 3
 AAW36230
 ID AAW36230 standard; Protein; 615 AA.
 XX
 AC AAW36230;
 XX
 DT 25-MAR-2003 (updated)
 DT 23-JUN-1995 (first entry)

```
XX DE Mycoplasma gallisepticum antigen (UM-67).
XX KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX OS Mycoplasma gallisepticum.
XX FH Key Location/Qualifiers
XX FT Protein 1..615
XX FT /note= "Trp residues correspond to TGA codons"
XX PN WO9423019-Al.
XX PD 13-OCT-1994.
XX PF 31-MAR-1994; 94WO-JP00541.
XX PR 31-MAR-1993; 93JP-0074139.
XX PR 30-SEP-1993; 93JP-0245625.
XX PA (JAPG ) NIPPON ZEON KK.
XX PA (SHIO ) SHIONOGI & CO LTD.
XX PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX PI Saeki S, Saitos, Takahashi K;
XX DR WPI; 1994-333181/41.
XX DR N-PSDB; AAQ77857.
XX PT Recombinant avipox virus combining DNA encoding a polypeptide -
XX PT exhibiting antigenicity of mycoplasma, useful for the production
XX PT of a live vaccine
XX PS Claim 4; Page 87-91; 123pp; Japanese.
XX CC A restriction fragment of the insert of M.gallisepticum genomic clone
XX CC PUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX CC ORF encodes an antigenic polypeptide (AAH63230). A recombinant avipox
XX CC virus comprising the coding sequence can be used as a live vaccine to
XX CC protect against infection by Mycoplasma gallisepticum.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 615 AA;
Query Match 95.3%; Score 1910; DB 15; Length 615;
Best Local Similarity 98.2%; Pred. No. 7e-122;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKARTLASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 122 STAYNQIRNRLVDLYNKASLLTKTLDPLNGTLLDSNEITTKANKNINNTLSTINEOKTN 181
DB 147 STAYNQIRNRLVDLYNKASLLTKTLDPLNGGMLDSNEITTVNRNINNTLSTINEOKTN 206
QY 182 ADALNSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
DB 207 ADALNSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 242 PSSRILANTNSITDVSIIWISLAGTNTKYQFSNSGTPSGYLYFFPKLYKAADANNVGLQ 301
DB 267 PSSRILANTNSITDVSIIWISLAGTNTKYQFSNSGTPSGYLYFFPKLYKAADANNVGLQ 326
QY 302 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 361
DB 327 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSSNENNADKI 384
```

```
DB 387 MNKVAPMIGNIYLSSNENNADKI 409
RESULT 4
AAR63229 standard; Protein; 610 AA.
XX ID AAR63229;
XX AC AAR63229;
XX XX
XX DT 25-MAR-2003 (updated)
XX DT 23-JUN-1995 (first entry)
XX XX
XX DE Mycoplasma gallisepticum antigen (UM-66).
XX KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX OS Mycoplasma gallisepticum.
XX FH Key Location/Qualifiers
XX FT Protein 1..610
XX FT /note= "Trp residues correspond to TGA codons"
XX XX
XX PN WO9423019-Al.
XX XX
XX PD 13-OCT-1994.
XX PF 31-MAR-1994; 94WO-JP00541.
XX PR 31-MAR-1993; 93JP-0074139.
XX PR 30-SEP-1993; 93JP-0245625.
XX XX
XX PA (JAPG ) NIPPON ZEON KK.
XX PA (SHIO ) SHIONOGI & CO LTD.
XX XX
XX PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX PI Saeki S, Saitos, Takahashi K;
XX DR WPI; 1994-333181/41.
XX DR N-PSDB; AAQ77856.
XX XX
XX PT Recombinant avipox virus combining DNA encoding a polypeptide -
XX PT exhibiting antigenicity of mycoplasma, useful for the production
XX PT of a live vaccine
XX PS Claim 4; Page 78-81; 123pp; Japanese.
XX CC A restriction fragment of the insert of M.gallisepticum genomic clone
XX CC PUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX CC ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
XX CC virus comprising the coding sequence can be used as a live vaccine to
XX CC protect against infection by Mycoplasma gallisepticum.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 610 AA;
Query Match 92.6%; Score 1856; DB 15; Length 610;
Best Local Similarity 95.5%; Pred. No. 3.3e-118;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
DB 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 146
QY 122 STAYNQIRNRLVDLYNKASLLTKTLDPLNGTLLDSNEITTKANKNINNTLSTINEOKTN 181
DB 147 STAYNQIRNRLVDLYNKASLLTKTLDPLNGGMLDSNEITTKANKNINNTLSTINEOKTN 206
QY 182 ADALNSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
```

Db	207	ADALANSFIKEVIQNKQSFVGMETNTNVQPSNFSFVAFSADVTPVNYKIARTVWNGDE	266
Qy	242	PSRILANTNSITDYSWIYSIAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNVGLQ	301
Db	267	PSRILANTNSITDYSWIYSISLGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADASNVLQ	326
Qy	302	YKLNGNNGVQVEFATSTSANNTTANPTPAVDKIVAKIVLSGLRFGONTIELSVPTGEGN	361
Db	327	YKLNGNNGVQVEFATSTSANNTTANPTPAVDKIVAKIVLSGLRFGONTIELSVPTGERN	386
Qy	362	MNKVAPMIGNIYLSNENNAADK 383	
Db	387	MNKVAPMIGNIYLSNENNAADK 408	
RESULT 5			
ID	AAAR44493		
ID	AAAR44493	standard; Protein; 368 AA.	
AC	AAAR44493;		
XX			
DT	25-MAR-2003	(updated)	
DT	16-JUN-1994	(first entry)	
XX			
DE		Mycoplasma gallisepticum 40kD antigen.	
XX			
KW		Vaccine; mycoplasma infection; poultry; fowl.	
XX			
OS		Mycoplasma gallisepticum.	
XX			
FH		Key Location/Qualifiers	
FT		Misc-difference 262	
FT		Misc-difference 283	
FT		/note= "corresponds to NNN codon in AAQ53419"	
FT		/note= "corresponds to NNN codon in AAQ53419"	
XX			
PN		WO9324646-A1.	
XX			
PD		09-DEC-1993.	
XX			
PF		28-MAY-1993; 93WO-JP00715.	
XX			
PR		29-MAY-1992; 92JP-0138819.	
XX			
PA		(JAPG) NIPPON ZEON KK.	
PA		(SHIO) SHIONOGI & CO LTD.	
XX			
PI		Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;	
XX			
DR		WPI; 1993-405837/50.	
DR		N-PSDB; AAQ53419.	
XX			
PT		Mycoplasma gallisepticum antigen and DNA coding for it - useful	
PT		for vaccination of fowl against mycoplasma infections	
XX			
PS		Claim 2; Page 23-26; 37pp; Japanese.	
XX			
CC		The sequence coding for the 40kDa antigen was obtained by PCR	
CC		amplification of M.gallisepticum genomic DNA. The antigen reacts	
CC		with Mycoplasma-immune or Mycoplasma-infected serum and can be used	
CC		as a vaccine to protect fowl from M.gallisepticum infection.	
CC		(Updated on 25-MAR-2003 to correct PN field.)	
XX			
SQ		Sequence 368 AA;	
		Query Match 80.4%; Score 1612; DB 14; Length 368;	
		Best Local Similarity 95.0%; Pred. No. 7.1e-102;	
		Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;	
Qy	2	CMSITKDDPNNGQTQLEAARMETDLINAKAMTSLAQDYAKIEASLSAYSEATVN	61
Db	27	CMSITKDDPNNGQTQLEAARMETDLINAKAMTSLAQDYAKIEASLSAYSEATVN	86

CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle disease virus haemagglutinin neuraminidase and FpV
 CC sequences. The recombinant avipox virus is useful as a live vaccine
 CC to protect against infection by Mycoplasma gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 CC
 XX Sequence 368 AA;
 SQ

Query Match 80.4%; Score 1612; DB 15; Length 368;
 Best Local Similarity 95.0%; Pred. No. 7.1e-102;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 86
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 QY 122 STAYNQIRNLVDLYNKASSLITKTLDPNGGTLDSNEITTKANKNINNTLSTINEQKTN 181
 DB 147 STAYNQIRNLVDLYNKASSLITKTLDPNGGTLDSNEITTKANKNINNTLSTINEQKTN 206
 QY 182 ADALNSFIKKVIQNNQEQSFVCTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
 DB 207 ADALNSFIKKVIQNNQEQSFVCTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
 QY 242 PSSRILANTNSITDVSXIYSLAGTNTKYQFSFNYPSTGTYLYFPYKLVKAADANNVGLQ 301
 DB 267 PSSRILANTNSITDVSXIYSLAGTNTKYQFSFNYPSTGTYLYFPYKLVKAADANNVGLQ 326
 QY 302 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAK 338
 DB 327 YKLNGNVOQVEFATSTANNTTANPTQOLMLRLKLLK 363

RESULT 7
 AAR76955
 ID AAR76955 standard; Protein; 368 AA.
 XX
 AC AAR76955;
 XX
 XX 25-MAR-2003 (updated)
 DT 12-MAR-1996 (first entry)
 XX
 XX Mycoplasma gallisepticum antigenic protein TTM-1.
 DE
 XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.
 KW
 XX Mycoplasma gallisepticum.
 OS
 XX Key Location/Qualifiers
 FH
 FT Misc-difference 262
 FT /note= "any amino acid"
 FT Misc-difference 283
 FT /note= "any amino acid"
 XX
 PN JP07133295-A.
 XX
 XX 23-MAY-1995.
 PD
 XX 27-AUG-1993; 93JP-0213102.
 PF
 XX 27-AUG-1993; 93JP-0213102.
 PR
 XX (JAPC) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX WPI; 1995-220782/29.
 DR N-PSDB; AAO94711.
 XX

PT A new antigenic protein which reacts with Mycoplasma gallisepticum -
 PT is useful in a component vaccine for use against poultry infected
 PT with M. gallisepticum.
 XX
 PS Claim 6; Figs 5-6; 33pp; Japanese.
 XX
 CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
 CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
 CC infectious diseases in poultry, and as a diagnostic agent for
 CC M. gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 CC
 XX Sequence 368 AA;
 SQ

Query Match 80.0%; Score 1604; DB 16; Length 368;
 Best Local Similarity 94.4%; Pred. No. 2.5e-101;
 Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 2 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 61
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN 86
 QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 QY 122 STAYNQIRNLVDLYNKASSLITKTLDPNGGTLDSNEITTKANKNINNTLSTINEQKTN 181
 DB 147 STAYNQIRNLVDLYNKASSLITKTLDPNGGTLDSNEITTKANKNINNTLSTINEQKTN 206
 QY 182 ADALNSFIKKVIQNNQEQSFVCTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 241
 DB 207 ADALNSFIKKVIQNNQEQSFVCTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
 QY 242 PSSRILANTNSITDVSXIYSLAGTNTKYQFSFNYPSTGTYLYFPYKLVKAADANNVGLQ 301
 DB 267 PSSRILANTNSITDVSXIYSLAGTNTKYQFSFNYPSTGTYLYFPYKLVKAADANNVGLQ 326
 QY 302 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAK 338
 DB 327 YKLNGNVOQVEFATSTANNTTANPTQOLMLRLKLLK 363

RESULT 8
 AAP93646
 ID AAP93646 standard; protein; 235 AA.
 XX
 AC AAP93646;
 XX
 XX 25-MAR-2003 (updated)
 DT 11-MAY-1990 (first entry)
 XX
 XX Amino acid sequence of Mycoplasma gallisepticum (MGI) polypeptide.
 DE
 XX Mycoplasma gallisepticum; Poultry vaccine; ss;
 KW
 XX Mycoplasma gallisepticum.
 OS
 XX
 PN EP345021-A.
 XX
 PD 06-DEC-1989.
 XX
 XX 31-MAY-1989; 89EP-0305441.
 PF 31-MAY-1989; 89EP-0305441.
 XX
 XX 02-JUN-1988; 88JP-0136343.
 PR
 XX (JAPC) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI SEIYAKU KK.
 PA (JAPC) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI SEIYAKU KK.
 XX
 XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
 PI

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
XX WPI; 1989-358393/49.
DR N-PSDB; AAN92566.
XX
XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT vaccines.
XX
XX Disclosure; Fig.1a; 31pp; English.
PS
XX This amino acid sequence of MG1 is encoded by M1 DNA and elicits an
CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 235 AA;
Query Match 57.0%; Score 1142; DB 10; Length 235;
Best Local Similarity 97.0%; Pred. No. 4.1e-70;
Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEATVN 61
DB 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEATVN 60
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVYAKALKTTLEQRATNLEGLS 121
DB 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVYAKALKTTLEQRATNLEGLA 120
QY 122 STAYNOIRNNLVLYNNKASSLITKTLDPLNGGTLDSNEITVANKNNINNTLSTINEQKTN 181
DB 121 STAYNOIRNNLVLYNNKASSLITKTLDPLNGGTLDSNEITVANKNNINNTLSTINEQKTN 180
QY 182 ADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 236
DB 181 ADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 235
RESULT 9
AAR05081
ID AAR05081 standard; protein; 235 AA.
XX AC AAR05081;
XX
XX 10-MAR-2003 (updated)
DT 08-OCT-1990 (first entry)
XX
XX MG-1 antigen.
XX
XX Mycoplasma gallisepticum; poultry; vaccine.
XX
XX Mycoplasma gallisepticum.
OS
XX
XX JP02111795-A.
PN
XX
XX 24-APR-1990.
PD
XX
XX 02-JUN-1989; 89JP-0136343.
PF
XX
XX 02-JUN-1989; 89JP-0136343.
PR
XX
XX (JAPG) NIPPON ZEON KK.
PA (SHIO) SHIONOGI KK.
XX
XX WPI; 1990-169109/22.
DR N-PSDB; AAQ04686.
XX
XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
PT antigen protein of the disease and recombinant vector
PT incorporated with its coding gene.

XX Claim 2; Fig 1a; 20pp; Japanese.
PS
XX DNA encoding the protein can be inserted into an expression vector
CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
CC be ligated to other DNA to produce fusion proteins with an N-terminal
CC bacterial enzyme sequence.
CC See also AAR05081-2 and AAR06437-41.
CC (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ Sequence 235 AA;
Query Match 57.0%; Score 1142; DB 11; Length 235;
Best Local Similarity 97.0%; Pred. No. 4.1e-70;
Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEATVN 61
DB 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEATVN 60
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVYAKALKTTLEQRATNLEGLS 121
DB 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHNPVYAKALKTTLEQRATNLEGLA 120
QY 122 STAYNOIRNNLVLYNNKASSLITKTLDPLNGGTLDSNEITVANKNNINNTLSTINEQKTN 181
DB 121 STAYNOIRNNLVLYNNKASSLITKTLDPLNGGTLDSNEITVANKNNINNTLSTINEQKTN 180
QY 182 ADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 236
DB 181 ADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 235
RESULT 10
AAR05082
ID AAR05082 standard; protein; 261 AA.
XX AC AAR05082;
XX
XX 10-MAR-2003 (updated)
DT 08-OCT-1990 (first entry)
XX
XX TMG-1 antigen.
XX
XX Mycoplasma gallisepticum; poultry; vaccine.
XX
XX Mycoplasma gallisepticum.
OS
XX
XX JP02111795-A.
PN
XX
XX 24-APR-1990.
PD
XX
XX 02-JUN-1989; 89JP-0136343.
PF
XX
XX 02-JUN-1989; 89JP-0136343.
PR
XX
XX (JAPG) NIPPON ZEON KK.
PA (SHIO) SHIONOGI KK.
XX
XX WPI; 1990-169109/22.
DR N-PSDB; AAQ04687.
XX
XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
PT antigen protein of the disease and recombinant vector
PT incorporated with its coding gene.
XX
XX Claim 2; Fig 2; 20pp; Japanese.
PS
XX DNA encoding the protein can be inserted into an expression vector
CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
CC be ligated to other DNA to produce fusion proteins with an N-terminal

CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 XX SQ Sequence 261 AA;
 Query Match 57.0%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 4.7e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 61
 DB 27 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 86
 QY 62 NNLNATLEOLKMAKTNLESAIQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
 DB 87 NNLNATLEOLKMAKTNLESAIQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLA 146
 QY 122 STAYNOIRNLVDLYNKASSLITKTLDPNGGMLDSNEITTNANKNINNTLSTINEQKTN 181
 DB 147 STAYNOIRNLVDLYNKASSLITKTLDPNGGMLDSNEITTNANKNINNTLSTINEQKTN 206
 QY 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSEAFVSADVTPVNYKYARTV 236
 DB 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSEAFVSADVTPVNYKYARTV 261
 RESULT 11
 AAR79911
 ID AAR79911 standard; Protein; 261 AA.
 XX AC AAR79911;
 XX DT 19-JUL-1996 (first entry)
 XX DE M.gallisepticum 261 amino acid protein.
 XX KW Detection; probe; primer; PCR; amplification; secretion; lung;
 XX KW avian chronic respiratory disease; respiratory tract; nasal cavity.
 XX OS Mycoplasma gallisepticum.
 XX PN JP07236498-A.
 XX PD 12-SEP-1995.
 XX PF 25-FEB-1994; 94JP-0052764.
 XX PR 25-FEB-1994; 94JP-0052764.
 XX PA (JAPG) NIPPON ZEON KK.
 XX PA (SHIO) SHIONOGI & CO LTD.
 XX WPI; 1995-347462/45.
 XX DR N-PSDB; AAT04076.
 XX PT Detection of Mycoplasma gallisepticum - for the quick detection,
 XX PT i.e. within one day, of avian chronic respiratory diseases
 XX PS Claim 3; Page 10-11; 11pp; Japanese.
 XX This is the amino acid sequence of a 261 amino acid protein encoded
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid
 CC protein) can be used to detect M.gallisepticum using probes based on
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
 CC complement of bases 893-919, 1908-1934 and the complement of bases
 CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
 CC encoding this protein. The method using these sequences is faster i.e. is
 CC able to detect M.gallisepticum, which causes avian chronic respiratory
 CC diseases, within one day, from avian secretions, washings from the lung,
 CC respiratory tract, nasal cavity, etc.

SQ Sequence 261 AA;
 Query Match 57.0%; Score 1142; DB 16; Length 261;
 Best Local Similarity 97.0%; Pred. No. 4.7e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 61
 DB 27 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 86
 QY 62 NNLNATLEOLKMAKTNLESAIQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 121
 DB 87 NNLNATLEOLKMAKTNLESAIQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLA 146
 QY 122 STAYNOIRNLVDLYNKASSLITKTLDPNGGMLDSNEITTNANKNINNTLSTINEQKTN 181
 DB 147 STAYNOIRNLVDLYNKASSLITKTLDPNGGMLDSNEITTNANKNINNTLSTINEQKTN 206
 QY 182 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSEAFVSADVTPVNYKYARTV 236
 DB 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSEAFVSADVTPVNYKYARTV 261
 RESULT 12
 AAP93959
 ID AAP93959 standard; protein; 261 AA.
 XX AC AAP93959;
 XX DT 25-MAR-2003 (updated)
 XX DT 11-MAY-1990 (first entry)
 XX DE Amino acid (AA) sequence of TMG-1 polypeptide.
 XX KW Mycoplasma gallisepticum; Poultry vaccine; ss;
 XX OS Mycoplasma gallisepticum.
 XX PN EP345021-A.
 XX PD 06-DEC-1989.
 XX PF 31-MAY-1989; 89EP-0305441.
 XX PF 31-MAY-1989; 89EP-0305441.
 XX PR 02-JUN-1988; 88JP-0136343.
 XX PA (JAPG) NIPPON ZEON KK.
 XX PA (SHIO) SHIONOGI SEIYAKU KK.
 XX PA (JAPG) NIPPON ZEON KK.
 XX PA (SHIO) SHIONOGI SEIYAKU KK.
 XX PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
 XX PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
 XX DR WPI; 1989-358393/49.
 XX DR N-PSDB; AAN92574.
 XX PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
 XX PT vaccines.
 XX PS Disclosure; Fig.2; 31pp; English.
 XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
 CC same sequence as that of a polypeptide expressed in Mycoplasma
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted
 CC into a recombinant vector used to transform a host the antigen protein
 CC produced can be used as a vaccine to prevent a host the antigen protein.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX SQ Sequence 261 AA;

Query Match 55.7%; Score 1117; DB 10; Length 261;
Best Local Similarity 96.2%; Pred. No. 2.4e-68;
Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;
QY 2 CMSTTKDANPNNGQTQLEARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 61
DB 27 CMSTTKDANPNNGQTQLEARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLA 146
QY 122 STAYNQIRNLVDLYNKASSLTKTLDPLNGGTLDSNEITTKANKNNLTSTINE-OKT 180
DB 147 STAYNQIRNLVDLYNKASSLTKTLDPLNGGTLDSNEITTKANKNNLTSTINE-OQT 205
QY 181 NADALSNSFKKVIQNNQEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 236
DB 206 NADALSNSFKKVIQNNQEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 261

RESULT 13
AAR63226 standard; Protein; 661 AA.
XX AAR63226;
XX
XX
DT 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
DE Mycoplasma gallisepticum antigen (UM-81).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
OS Mycoplasma gallisepticum.
FH Key Location/Qualifiers
FT Protein 1..661
FT /note= "Tripeptides correspond to TGA codons"

XX W09423019-A1.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94MO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Sasaki S, Saitos, Takahashi K;
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77853.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 61-65; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-81 containing an open reading frame was sequenced (AAQ77853). The
XX ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 661 AA;
XX

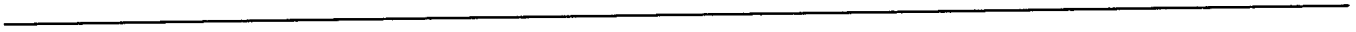
Query Match 40.2%; Score 806; DB 15; Length 661;
Best Local Similarity 42.8%; Pred. No. 1.3e-46;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;
QY 9 DANPNNGQTQLEARMELTD-----LINAKAMTILASLDQYAKIEASLSAYSEATVN 61
DB 45 NTPNSDGGGMMNAAKAKELADAKAALTTTLINGETANLASIEDYAKIKSELTSAIETKAVS 104
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121
DB 105 AKTGATLEVNFAKTTLDAAIKAKASAKNDFDAQHGSLVEAYNNLKETLKEKTNLSLA 164
QY 122 STAYNQIRNLVDLYNKASSLTKTLDPLNGGTLDSNEITTKANKNNLTSTINEOQTN 181
DB 165 NENYAAIRTNLSLYEKANTIVTATLDPAT-GNIEPVMSVTQANQDITNATSRLLIAWKON 223
QY 182 ADALSNSFKKVIQNNQEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGE 241
DB 224 ADNLANSFKKVIQNNQEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGE 282
QY 242 -----PSSRILANTNSITDVSNIYSLAGTNTKYQSFNSY- PSTGYLYFFYKLVKKAAD 294
DB 283 TPLATTPAEDATQQAASLTDVSNWISLNGAEAKYTLSEFYFGAEKATYLYFFYKLVKTS 342
QY 295 ANNVLQYKLNNGVQOQVEF-----ATSTAN--NTTANP-----TPAVDEIKVAKIV 340
DB 343 --NVGLQYKLNNGVQOQVEF-----ATSTAN--NTTANP-----TPAVDEIKVAKIV 400
QY 341 LSGLRFGQNTIELSVPTGEGNMKNKAPMIGNIYLS-----NENN 380
DB 401 LSNLKNFNSNTEFVSPTG-----KAAPMIGNIYLSNENKNIYDDLFGNSFNENN 455
QY 381 -----ADKIPGY 387
DB 456 PTATVVDLLKGY 467
RESULT 14
AAR79910 standard; Protein; 661 AA.
XX AAR79910;
XX
XX 19-JUL-1996 (first entry)
XX
XX M.gallisepticum 661 amino acid protein.
XX
XX Detection; probe; primer; PCR; amplification; secretion; lung;
XX avian chronic respiratory disease; respiratory tract; nasal cavity.
XX
XX Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
FH Misc-difference 220 /note= "encoded by TGA"
FT Misc-difference 270 /note= "encoded by TGA"
FT Misc-difference 305 /note= "encoded by TGA"
FT Misc-difference 581 /note= "encoded by TGA"
XX
XX JP07236498-A.
XX
XX 12-SEP-1995.
XX
XX 25-FEB-1994; 94JP-0052764.
XX
XX 25-FEB-1994; 94JP-0052764.
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX

[illegible]

us-09-147-052-4_copy_693_1086.rag

Fri Oct 10 11:43:21 2003

Search completed: October 8, 2003, 17:29:31
Job time : 84 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:58:45 ; Search time 57.8445 Seconds
(without alignments)
2997.887 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score: 3510

Sequence: 1 MHYFRNCIFFLIVILGTN.....DVGVDYAEVARRNQLHELK 672

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3510	100.0	805	12	Q8Y44 turkey herp
2	3507	99.9	865	12	Q8JLW4 turkey herp
3	3503	99.8	805	12	Q8Y45 turkey herp
4	3503	99.8	865	12	Q8JLW2 turkey herp
5	3496	99.6	865	12	Q8JLW3 turkey herp
6	3494	99.5	865	12	Q8JLW3 turkey herp
7	2990.5	85.2	865	12	Q8JLW3 turkey herp
8	2975.5	84.8	865	12	Q8JLW3 turkey herp
9	2971	84.6	870	12	Q8JLW3 turkey herp
10	2967	84.5	864	12	Q8JLW3 turkey herp
11	2967	84.5	870	12	Q8JLW3 turkey herp
12	1895.5	54.0	879	12	Q8JLW3 turkey herp
13	1868	53.2	881	12	Q8JLW3 turkey herp
14	1850	52.7	948	12	Q8JLW3 turkey herp
15	1754.5	50.0	919	12	Q8JLW3 turkey herp
16	1748.5	49.8	931	12	Q8JLW3 turkey herp

17	1735.5	49.4	980	12	Q66682 equine herp
18	1734	49.4	868	12	Q8Y44 human herp
19	1721.5	49.0	950	12	Q66678 equine herp
20	1707	48.6	891	12	Q9J052 simian herp
21	1703.5	48.5	906	12	Q9J052 simian herp
22	1702.5	48.5	975	12	Q9J052 simian herp
23	1700	48.4	891	12	Q66018 simian herp
24	1698	48.4	920	12	Q8Y44 human herp
25	1696.5	48.3	947	12	Q9J052 simian herp
26	1696	48.3	904	12	Q9J052 simian herp
27	1696	48.3	912	12	Q8Y44 human herp
28	1695	48.3	904	12	Q8Y44 human herp
29	1694	48.3	901	12	Q8Y44 human herp
30	1692	48.2	904	12	Q8Y44 human herp
31	1691	48.2	901	12	Q8Y44 human herp
32	1690	48.1	929	12	Q8Y44 human herp
33	1683.5	48.0	943	12	Q8Y44 human herp
34	1679	47.8	885	12	Q8Y44 human herp
35	1678	47.8	902	12	Q8Y44 human herp
36	1672	47.6	908	12	Q8Y44 human herp
37	1667	47.5	893	12	Q8Y44 human herp
38	1664	47.4	904	12	Q8Y44 human herp
39	1663	47.4	904	12	Q8Y44 human herp
40	1660	47.3	903	12	Q8Y44 human herp
41	1660	47.3	904	12	Q8Y44 human herp
42	1656	47.2	904	12	Q8Y44 human herp
43	1647.5	46.9	916	12	Q8Y44 human herp
44	1630.5	46.5	894	12	Q8Y44 human herp
45	1619.5	46.1	887	12	Q8Y44 human herp

ALIGNMENTS

RESULT 1

Q8Y44 PRELIMINARY; PRT; 805 AA.

AC Q8Y44; TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein B (Fragment).
OS Turkey herpesvirus.
OC Viruses; GSDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson I.;
RT "Marek's disease virus - turkey isolate, gb.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350326; AAK37557.2;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
FT NON_TER 805
SQ SEQUENCE 805 AA; 91115 MW; FC8EB090885E15AA CRC64;

Query Match 100.0%; Score 3510; DB 12; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.7e-270; Indels 0; Gaps 0;
Matches 672; Conservative 0; Mismatches 0;

QY	1	MHYFRNCIFFLIVILYGTNSPSTQNTVSREWSSVOLSEESTFYLCPPVGSTVIRL	60
Db	1	MHYFRNCIFFLIVILYGTNSPSTQNTVSREWSSVOLSEESTFYLCPPVGSTVIRL	60
QY	61	EPKRCPEPKATENGEGTALFKENISPYFKVLYYKNIQTWTGTYRQITNRYT	120
Db	61	EPKRCPEPKATENGEGTALFKENISPYFKVLYYKNIQTWTGTYRQITNRYT	120
QY	121	DRTPVSIETDLDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPESRAW	180
Db	121	DRTPVSIETDLDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPESRAW	180

```

QY 181 HTTNETTYVWGSPWIIYRTGTSVNCIVEEMDARSFPYPSYFAMANGDIANISPPFYGLSPPE 240
Db 181 HTTNETTYVWGSPWIIYRTGTSVNCIVEEMDARSFPYPSYFAMANGDIANISPPFYGLSPPE 240
QY 241 AAAEPMGYPODNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGVWDWAPKTRVCSM 300
Db 241 AAAEPMGYPODNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGVWDWAPKTRVCSM 300
QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDNRILIGQCICKREAEAAIEQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDNRILIGQCICKREAEAAIEQIFR 360
QY 361 TKYNDSHVKVGHVOYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEMLDLVNNKHA 420
Db 361 TKYNDSHVKVGHVOYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEMLDLVNNKHA 420
QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVOPAMLOFLYDHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVOPAMLOFLYDHIQTHINDMFSR 480
QY 481 IATAWCELONRELVLVHHEGIKINPSATASATLGRVAAKMLGDVAASVCTAIDAESVTL 540
Db 481 IATAWCELONRELVLVHHEGIKINPSATASATLGRVAAKMLGDVAASVCTAIDAESVTL 540
QY 541 QNSMRVITSTNTCYSRPLVLFSGYNOGNIQOLGNNELLPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFSGYNOGNIQOLGNNELLPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKWDAADIIQASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
Db 601 GSGYALFENYFVKWDAADIIQASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
QY 661 EVARNQHLK 672
Db 661 EVARNQHLK 672

```

RESULT 2

Q8JLW4 PRELIMINARY; PRT; 865 AA.

ID Q8JLW4; 865 AA.

AC Q8JLW4; 865 AA.

DT 01-OCT-2002 (TremBLrel. 22, Created)

DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)

DE Glycoprotein B.

GN GB.

OS Turkey herpesvirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Marek's disease-like viruses.

OX NCBI_TaxID-10390;

RN [1]

RC SEQUENCE FROM N.A.

RA Parcells M.S., Shamblin C.E., Dienglewiecz R.L.;

RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three

RT Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations

RT in the Glycoprotein L-encoding Gene in Some vv+MDVs."

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY129965; AAM97698.1;

DR InterPro; IPR00234; Glycoprot.B.

DR Pfam; PF00606; Glycoprotein_B_1.

DR ProDom; PD000693; Glycoprot_B; 1.

SQ SEQUENCE 865 AA; 98107 MW; 4D2G28B5E4DEB28B CRC64;

```

Query Match 99.9%; Score 3507; DB 12; Length 865;
Best Local Similarity 99.9%; Pred. No. 3.2e-270;
Matches 671; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVILYGNSSPSTQNTVSREVSVQVLSSESTFYLCPPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVILYGNSSPSTQNTVSREVSVQVLSSESTFYLCPPPVGVSTVIRL 60

```

```

QY 61 EPPKRCPEPRKATEWEGIAILFKENISPYKFKVLYKNIIOQTWTGTTTQITNRYT 120
Db 61 EPPKRCPEPRKATEWEGIAILFKENISPYKFKVLYKNIIOQTWTGTTTQITNRYT 120
QY 121 DRTPVSEETDLDIDGKRCSSKARYLRNNVYVEAFDRDAGEKOVLLKPSKFNTPESRAW 180
Db 121 DRTPVSEETDLDIDGKRCSSKARYLRNNVYVEAFDRDAGEKOVLLKPSKFNTPESRAW 180
QY 181 HTTNETTYVWGSPWIIYRTGTSVNCIVEEMDARSFPYPSYFAMANGDIANISPPFYGLSPPE 240
Db 181 HTTNETTYVWGSPWIIYRTGTSVNCIVEEMDARSFPYPSYFAMANGDIANISPPFYGLSPPE 240
QY 241 AAAEPMGYPODNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGVWDWAPKTRVCSM 300
Db 241 AAAEPMGYPODNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGVWDWAPKTRVCSM 300
QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDNRILIGQCICKREAEAAIEQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDNRILIGQCICKREAEAAIEQIFR 360
QY 361 TKYNDSHVKVGHVOYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEMLDLVNNKHA 420
Db 361 TKYNDSHVKVGHVOYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEMLDLVNNKHA 420
QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVOPAMLOFLYDHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVOPAMLOFLYDHIQTHINDMFSR 480
QY 481 IATAWCELONRELVLVHHEGIKINPSATASATLGRVAAKMLGDVAASVCTAIDAESVTL 540
Db 481 IATAWCELONRELVLVHHEGIKINPSATASATLGRVAAKMLGDVAASVCTAIDAESVTL 540
QY 541 QNSMRVITSTNTCYSRPLVLFSGYNOGNIQOLGNNELLPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFSGYNOGNIQOLGNNELLPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKWDAADIIQASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
Db 601 GSGYALFENYFVKWDAADIIQASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
QY 661 EVARNQHLK 672
Db 661 EVARNQHLK 672

```

RESULT 3

Q98Y45 PRELIMINARY; PRT; 805 AA.

ID Q98Y45; 805 AA.

AC Q98Y45; 805 AA.

DT 01-JUN-2001 (TremBLrel. 17, Created)

DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)

DE Glycoprotein B (Fragment).

OS Turkey herpesvirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Marek's disease-like viruses.

OX NCBI_TaxID-10390;

RN [1]

RP SEQUENCE FROM N.A.

RA Davidson I.;

RT "Marek's disease virus - chicken isolate, gb gene."

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF350325; AAK37556.2;

DR InterPro; IPR000234; Glycoprot.B.

DR Pfam; PF00606; Glycoprotein_B; 1.

DR ProDom; PD000693; Glycoprot_B; 1.

FT NON_TER 805 805

SQ SEQUENCE 805 AA; 91111 MW; F88B4410A85E15AA CRC64;

```

Query Match 99.8%; Score 3503; DB 12; Length 805;
Best Local Similarity 99.9%; Pred. No. 6e-270;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MHVFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVQLSEESTFYLCPPVGVSTVIRL 60
 DB 1 MHVFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVQLSEESTFYLCPPVGVSTVIRL 60
 QY 61 EPPRCPEPRKATENGEGIAILFKENISPYKFKVLYYKNIQTWTGTTTYRQITNRYT 120
 DB 61 EPPRCPEPRKATENGEGIAILFKENISPYKFKVLYYKNIQTWTGTTTYRQITNRYT 120
 QY 121 DRTPSVIEITLDIGKGRCSKARYLRNNVYVEAFDRDAGEQVLLKPKSKFNTPESSRAW 180
 DB 121 DRTPSVIEITLDIGKGRCSKARYLRNNVYVEAFDRDAGEQVLLKPKSKFNTPESSRAW 180
 QY 181 HTTNETYTVWGSFWIYRTGTSVNCIYEEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
 DB 181 HTTNETYTVWGSFWIYRTGTSVNCIYEEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
 QY 241 AAAEPMGYPQDNFKQLDSYFSDLDKRRKASLPVKRNFLLTSHFTVGVDMWAPKTRVCSM 300
 DB 241 AAAEPMGYPQDNFKQLDSYFSDLDKRRKASLPVKRNFLLTSHFTVGVDMWAPKTRVCSM 300
 QY 301 TKWKEVTEMLRATVNGRYRFMARLSATFISNTTFDPNRIILGQCICKREAAAEIQIFR 360
 DB 301 TKWKEVTEMLRATVNGRYRFMARLSATFISNTTFDPNRIILGQCICKREAAAEIQIFR 360
 QY 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYRLMLRDNRTDEMLDLVNKHHA 420
 DB 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYRLMLRDNRTDEMLDLVNKHHA 420
 QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVOFAMQLFYLDHIQTHINDMFSR 480
 DB 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVOFAMQLFYLDHIQTHINDMFSR 480
 QY 481 IATANCELQNRRELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
 DB 481 IATANCELQNRRELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELLPVYVTKKEELRDVGVLDTA 660
 DB 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELLPVYVTKKEELRDVGVLDTA 660
 QY 661 EVARRNQLHELK 672
 DB 661 EVARRNQLHELK 672

RESULT 4

Q8JLW2 PRELIMINARY; PRT; 865 AA.
 ID Q8JLW2
 AC Q8JLW2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=attenuated GA;
 RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
 RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 RT Pathotypes of Marek's Disease Viruses (vMDV, vMDV, vvMDV): Mutations
 RL in the Glycoprotein 1-encoding Gene in Some vvMDVs.";
 DR EMBL; AV129969; AAM97702.1; -;
 DR InterPro: IPR000234; Glycoprot. B.
 DR Pfam; PF00606; Glycoprotein_B; 1.

DR

ProDom: PD000693; Glycoprot_B; 1.

SQ SEQUENCE 865 AA; 98020 MW; F4P526114EA02DE5 CRC64;

Query Match 99.8%; Score 3503; DB 12; Length 865;

Best Local Similarity 99.9%; Pred. No. 6.7e-270;

Matches 671; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 MHVFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVQLSEESTFYLCPPVGVSTVIRL 60
 DB 1 MHVFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVQLSEESTFYLCPPVGVSTVIRL 60
 QY 61 EPPRCPEPRKATENGEGIAILFKENISPYKFKVLYYKNIQTWTGTTTYRQITNRYT 120
 DB 61 EPPRCPEPRKATENGEGIAILFKENISPYKFKVLYYKNIQTWTGTTTYRQITNRYT 120
 QY 121 DRTPSVIEITLDIGKGRCSKARYLRNNVYVEAFDRDAGEQVLLKPKSKFNTPESSRAW 180
 DB 121 DRTPSVIEITLDIGKGRCSKARYLRNNVYVEAFDRDAGEQVLLKPKSKFNTPESSRAW 180
 QY 181 HTTNETYTVWGSFWIYRTGTSVNCIYEEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
 DB 181 HTTNETYTVWGSFWIYRTGTSVNCIYEEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
 QY 241 AAAEPMGYPQDNFKQLDSYFSDLDKRRKASLPVKRNFLLTSHFTVGVDMWAPKTRVCSM 300
 DB 241 AAAEPMGYPQDNFKQLDSYFSDLDKRRKASLPVKRNFLLTSHFTVGVDMWAPKTRVCSM 300
 QY 301 TKWKEVTEMLRATVNGRYRFMARLSATFISNTTFDPNRIILGQCICKREAAAEIQIFR 360
 DB 301 TKWKEVTEMLRATVNGRYRFMARLSATFISNTTFDPNRIILGQCICKREAAAEIQIFR 360
 QY 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYRLMLRDNRTDEMLDLVNKHHA 420
 DB 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYRLMLRDNRTDEMLDLVNKHHA 420
 QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVOFAMQLFYLDHIQTHINDMFSR 480
 DB 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVOFAMQLFYLDHIQTHINDMFSR 480
 QY 481 IATANCELQNRRELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
 DB 481 IATANCELQNRRELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELLPVYVTKKEELRDVGVLDTA 660
 DB 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELLPVYVTKKEELRDVGVLDTA 660
 QY 601 GSGYALFENYFVKWDAADIQIASTFVELNLTLLDREILPLSVYVTKKEELRDVGVLDTA 660
 DB 601 GSGYALFENYFVKWDAADIQIASTFVELNLTLLDREILPLSVYVTKKEELRDVGVLDTA 660
 QY 661 EVARRNQLHELK 672
 DB 661 EVARRNQLHELK 672

RESULT 5

Q8JLW3 PRELIMINARY; PRT; 865 AA.
 ID Q8JLW3
 AC Q8JLW3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Glycoprotein B.
 GN GB.
 OS Turkey herpesvirus
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N.
 RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;

"DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
 RT pathotypes of Marek's disease viruses (vMDV, vVMDV, vVMDV): Mutations
 in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129967; AAM97700.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 865 AA; 98030 MW; 9008958B4B4E114D CRC64;

Query Match 99.6%; Score 3496; DB 12; Length 865;
 Best Local Similarity 99.7%; Pred. No. 2.4e-269;
 Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVOLSEESTFYLCPPVGVSTVRL 60
 DB 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVOLSEESTFYLCPPVGVSTVRL 60
 QY 61 EPPKRCPEPRKATEWGEGIALFKENISPYKFKVTLTKYKNIQTTWTGTYRQITNRYT 120
 DB 61 EPPKRCPEPRKATEWGEGIALFKENISPYKFKVTLTKYKNIQTTWTGTYRQITNRYT 120
 QY 121 DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 DB 121 DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 QY 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE 240
 DB 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE 240
 QY 241 AAEPMPGYPQDNFKQLDSYFSMDLDRKRSKASLPVKRNFLITSHFTVGMWAPKTRVCSM 300
 DB 241 AAEPMPGYPQDNFKQLDSYFSMDLDRKRSKASLPVKRNFLITSHFTVGMWAPKTRVCSM 300
 QY 301 TKWEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCIRKREAAIEQIFR 360
 DB 301 TKWEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCIRKREAAIEQIFR 360
 QY 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMIDLNNKHA 420
 DB 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMIDLNNKHA 420
 QY 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLDYDHIQTHINDMFSR 480
 DB 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLDYDHIQTHINDMFSR 480
 QY 481 IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
 DB 481 IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLFSGYGENOGNIQGGQGENNELPTLEAVEPCSANHRRYFLF 600
 DB 541 QNSMRVITSTNTCYSRPLVLFSGYGENOGNIQGGQGENNELPTLEAVEPCSANHRRYFLF 600
 QY 601 GSGYALFENYFKVWDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVDYA 660
 DB 601 GSGYALFENYFKVWDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVDYA 660
 QY 661 EVARRNQLHELK 672
 DB 661 EVARRNQLHELK 672

RESULT 6

Q83291 PRELIMINARY; PRT; 865 AA.
 ID Q83291
 AC Q83291;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Glycoprotein 100 precursor.
 GN GB.
 OS Marek disease virus type 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=38013;
 RN [1]
 RP SEQUENCE OF 1-498 FROM N.A.
 RC STRAIN=JM;
 RA Souloparov M.A., Bakhtina M.M., Krendelshchikov A.V., Babkin I.V.;
 RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
 RL complex gp100, gp60, gp48 (b-antigen).";
 RL submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X91985; CAA63039.1; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Signal.
 FT SIGNAL
 SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;

Query Match 99.5%; Score 3494; DB 12; Length 865;
 Best Local Similarity 99.7%; Pred. No. 3.5e-269;
 Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVOLSEESTFYLCPPVGVSTVRL 60
 DB 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVOLSEESTFYLCPPVGVSTVRL 60
 QY 61 EPPKRCPEPRKATEWGEGIALFKENISPYKFKVTLTKYKNIQTTWTGTYRQITNRYT 120
 DB 61 EPPKRCPEPRKATEWGEGIALFKENISPYKFKVTLTKYKNIQTTWTGTYRQITNRYT 120
 QY 121 DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 DB 121 DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 QY 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE 240
 DB 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE 240
 QY 241 AAEPMPGYPQDNFKQLDSYFSMDLDRKRSKASLPVKRNFLITSHFTVGMWAPKTRVCSM 300
 DB 241 AAEPMPGYPQDNFKQLDSYFSMDLDRKRSKASLPVKRNFLITSHFTVGMWAPKTRVCSM 300
 QY 301 TKWEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCIRKREAAIEQIFR 360
 DB 301 TKWEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCIRKREAAIEQIFR 360
 QY 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMIDLNNKHA 420
 DB 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMIDLNNKHA 420
 QY 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLDYDHIQTHINDMFSR 480
 DB 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLDYDHIQTHINDMFSR 480
 QY 481 IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
 DB 481 IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLFSGYGENOGNIQGGQGENNELPTLEAVEPCSANHRRYFLF 600
 DB 541 QNSMRVITSTNTCYSRPLVLFSGYGENOGNIQGGQGENNELPTLEAVEPCSANHRRYFLF 600
 QY 601 GSGYALFENYFKVWDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVDYA 660
 DB 601 GSGYALFENYFKVWDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVDYA 660
 QY 661 EVARRNQLHELK 672
 DB 661 EVARRNQLHELK 672

RESULT 7
Q9PWZ1


```

ID Q9PWZ1 PRELIMINARY; PRT; 865 AA.
AC Q9PWZ1.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
DE Homolog of HSV-1 major DNA-binding protein (Glycoprotein B).
GN ORF 35 OR UL27.
OS Turkey herpesvirus,
OS Marek's disease virus serotype 2 MDV2, and
OS Gallid herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID-10390, 36353, 35250;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-turkey herpesvirus; STRAIN-HPRS24;
RA Kato K., Jang H., Izumiya Y., Cai J., Teushima Y., Miyazawa T.,
RA Kai C., Mikami T.;
RT "Identification and Transcriptional Analysis of the Marek's Disease
RT Virus Serotype 2 Genes Homologous to the Glycoprotein B (UL27), the
RT ICP18.5 (UL28) and the Major DNA-binding Protein (UL29) Genes of
RT Herpes Simplex Virus Type 1.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 1; STRAIN-HPRS24;
RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
RA Lee Y., Kai C., Takahashi E., Mikami T.;
RT "The complete DNA sequence and transcription map of the unique long
RT genome region of Marek's disease virus type 2.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 3; STRAIN-HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 3; STRAIN-HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RT "A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
RT Strain HPRS24.";
RT Curr. Top. Microbiol. Immunol. 0:0-0(2000).
DR EMBL; AB024711; BAA83751.1; -.
DR EMBL; AB024414; BAA82923.1; -.
DR EMBL; AB049735; BAB16537.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR DNA-binding.
KW SEQUENCE 865 AA; 97309 MW; A20E66C70A475C35 CRC64;

Query Match 85.2%; Score 2990.5; DB 12; Length 865;
Best Local Similarity 83.1%; Pred. No. 4.3e-229;
Matches 559; Conservative 55; Mismatches 58; Indels 1; Gaps 1;

QY 1 MHYFRNCI-PFLIVILYXGNSPSTQNTVTSREVSVSVQSEESTFYLCPPPGVSTVIR 59
Db 1 MNHFGSICVPLFISVVFYFGRVARAQNVTGREAVSVSVQSEESTFYLCPPPGVSTVIR 60
QY 60 LEPPKPCPEPRKATWEGGIALFKENISPKYFKVLYKNIQTWTGTYRQITNRY 119
Db 61 LEPPKPCPEPRKATWEGGIALFKENINPYKFKVLYKNIQTWTGTYRQITNRF 120
QY 120 TDRTPVSEETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKOVLLKPSKFNTPESRA 179
Db 121 TDRTPVSEETDVIDAGRCSSKARYLRNNVYVEAYDGDAGEKOVLLQPSKFNTPESKA 180
QY 180 WHTTNETTVGSPWYRTGTSVNCIVEEMDARSVPFYSYFAMANGDIANTISPFYGLSP 239
Db 181 WHTTNDTVGSPWYRTGTSVNCIVEEMDARAAPFYSYFAMANGDIANNSPFYGVAPP 240
QY 240 EAAEPMGYPQDNFKQLDSYFSMDLDKRRKASLPVKRNFLLTSHTVGVGDWNAKTRVCS 299

||||| : : : : : ||| ||| : : : : : ||| : : : : : ||| : : : : : |||
241 EAAEPMGYPQDNFKQLDSYFSMDLSKROKASIPVKNRFLVTPHFTVGMDAAKTRVCS 300
QY 300 MTKKVEVTEMLRATVNGRYRFARELSATFISNTTFDPNRIILGQCICKREAAIEQIF 359
Db 301 MVKWKDVTEMLRATVNGRYRFARELSSTFISNTTFDPDRIRLGCQVKRDAETTKRIF 360
QY 360 RTKYNDSHVKGHVQYFLALGGFIVAYOPVLSKSLAHMYLRELMNDRTDMLDLVNNKH 419
Db 361 AQKYNDSHVKGHVQYFLALGGFIVAYOPVMSKSLAHMYLRELMNDRTDMLDLVNNKH 420
QY 420 AIYKKNATSLRLRRDIRNAPNRKTLDDTTAKTSSVQFAMQLQFLYDHIQTHNDMP 479
Db 421 ALSNKTIVSLRLRRDIRNAPNRKTLDDTTAKTSSVQFAMQLQFLYDHIQTHNDMP 480
QY 480 RIATAWCELQNRRLVHEGKINPISATATGLRRVAAKMLGDVAAVSSCTAIDAESVT 539
Db 481 RIATAWCELQNRRLVHEGKINPISATATGLRRVAAKMLGDVAAVSSCTAIDAESVT 540
QY 540 LQNSMRVITSTNTCYSRPLVLFSGYENOGNIOGQIGENNELPTLEAVEPCSANHRRYFL 599
Db 541 LQNSMRVATSTNMCYSRPLVLFSGYENOGRIQGLGENNELPTLEAVEPCSANHRRYFL 600
QY 600 FSGSYALFENYFVKWDAADIQIATSTFVELNLTLLEDEILPLSVYKKEELRDYGVLDY 659
Db 601 FGSAYALFEDYNFVAKVEVADIQTASTFVDLNLTLLEDEILPLSVYKKEELRDYGVLDY 660
QY 660 AEVARRNQLHELK 672
Db 661 AEVARRNQLHELK 673

RESULT 8
Q69406 PRELIMINARY; PRT; 865 AA.
ID Q69406
AC Q69406;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GB homolog.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID-10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SB-1;
RX MEDLINE-94233711; PubMed-8178437;
RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
RT Identification and expression by recombinant fowlpox viruses.";
RL Virology 200:484-493(1994).
DR EMBL; U01886; AAA19445.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR SEQUENCE 865 AA; 97236 MW; 04D4802FDC2E3948 CRC64;

Query Match 84.8%; Score 2975.5; DB 12; Length 865;
Best Local Similarity 82.8%; Pred. No. 6.7e-228;
Matches 557; Conservative 55; Mismatches 60; Indels 1; Gaps 1;

QY 1 MHYFRNCI-PFLIVILYXGNSPSTQNTVTSREVSVSVQSEESTFYLCPPPGVSTVIR 59
Db 1 MNHFGSICVPLFISVVFYFGRVARAQNVTGREAVSVSVQSEESTFYLCPPPGVSTVIR 60
QY 60 LEPPKPCPEPRKATWEGGIALFKENISPKYFKVLYKNIQTWTGTYRQITNRY 119
Db 61 LEPPKPCPEPRKATWEGGIALFKENINPYKFKVLYKNIQTWTGTYRQITNRF 120
QY 120 TDRTPVSEETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKOVLLKPSKFNTPESRA 179
Db 121 TDRTPVSEETDVIDAGRCSSKARYLRNNVYVEAYDGDAGEKOVLLQPSKFNTPESKA 180

```

QY 180 WHTTNTYTVWGSFWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFYGLSP 239
 DB 181 WHTTNTYTVWGSFWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFYGLSP 239
 QY 240 EAAAEPMGYPQDQNFKOLDYSFMDLDRKASLPVKNFLITSHFTVGDWMAKPTTRVCS 299
 DB 241 EAAAEPMGYPQDQNFKOLDYSFMDLDRKASLPVKNFLITSHFTVGDWMAKPTTRVCS 299
 QY 300 MTKWKEVTEMLRATVNGRYRFMAARELSATFISNTTEFDPNRIILGOCIKREAEAEQIF 359
 DB 301 MTKWKEVTEMLRATVNGRYRFMAARELSATFISNTTEFDPNRIILGOCIKREAEAEQIF 359
 QY 360 RRYKNDSHVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKH 419
 DB 361 AOKYNDSHVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKH 419
 QY 420 AIYKKNATSLSLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLQFLYDHIQTHINDMFS 479
 DB 421 AIYKKNATSLSLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLQFLYDHIQTHINDMFS 479
 QY 480 RIATAWCELQNLRELRLVWHEGIIKINPSATASATLGRVAAKMLGDVAASVSSCTAIDAESVT 539
 DB 481 RIATAWCELQNLRELRLVWHEGIIKINPSATASATLGRVAAKMLGDVAASVSSCTAIDAESVT 539
 QY 540 LQNSMRVITSTNTCYSRPLVLFYSYGENOGNIQOLGQENNELPTLEAVEPCSANHRRYFL 599
 DB 541 LQNSMRVITSTNTCYSRPLVLFYSYGENOGNIQOLGQENNELPTLEAVEPCSANHRRYFL 599
 QY 600 FGSYALFENYFVKMVAADIIQIASTFVELMLTLEDEILPLSVYTKKEELRDVGVLGY 659
 DB 601 FGSYALFENYFVKMVAADIIQIASTFVELMLTLEDEILPLSVYTKKEELRDVGVLGY 659
 QY 660 AEVARNOLHELK 672
 DB 661 AEVARNOLHELK 673

RESULT 9

QSDPQ9 ID Q9DPQ9 PRELIMINARY; PRT: 870 AA.
 AC Q9DPQ9; STRAIN=FC126;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE UL27 virion membrane glycoprotein B.
 GN HV7035.
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=37108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC126;
 RA Medline=20578232; PubMed=11134310;
 RX Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
 RT "The genome of turkey herpesvirus."
 RL J. Virol. 75:971-978(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC126;
 RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR ENBL: AF291866; AAC45765.1; --
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 870 AA; 98816 MW; F93D1C7036938FB1 CRC64;

Query Match 84.68; Score 2971; DB 12; Length 870;
 Best Local Similarity 82.08; Pred. No. 1.6e-227;
 Matches 551; Conservative 57; Mismatches 64; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLLVILYCTNSSPSTONVTSREVSVQVSEESTFYLCPPPVGSVIRL 60
 DB 7 MYFNKSLFELTLPILSIATSEIKLPNVTAREIVSGIQLSEDETTFVCPVPSIVRL 66
 QY 61 EPRKCEPEKATWEGEGAILFKENISPKFKVLYLYKNIQITTTWTGTTTQIQRNYT 120
 DB 67 EPRKCEPEKATWEGEGAILFKENISPKFKVLYLYKNIQITTTWTGTTTQIQRNYT 126
 QY 121 DRTPSIEITLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPESRAW 180
 DB 127 DRTPSIEITLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPESRAW 186
 QY 181 HTTNETYTVWGSFWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFYGLSPPE 240
 DB 187 HTTNETYTVWGSFWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFYGLSPPE 246
 QY 241 AAAEPMGYPQDQNFKOLDYSFMDLDRKASLPVKNFLITSHFTVGDWMAKPTTRVCSM 300
 DB 247 AAAEPMGYPQDQNFKOLDYSFMDLDRKASLPVKNFLITSHFTVGDWMAKPTTRVCSM 306
 QY 301 TWKKEVTEMLRATVNGRYRFMAARELSATFISNTTEFDPNRIILGOCIKREAEAEQIFR 360
 DB 307 TWKKEVTEMLRATVNGRYRFMAARELSATFISNTTEFDPNRIILGOCIKREAEAEQIFR 366
 QY 361 RYKNDSHVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKH 420
 DB 367 RYKNDSHVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKH 426
 QY 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLQFLYDHIQTHINDMFS 480
 DB 427 IYKKNATSLSLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLQFLYDHIQTHINDMFS 486
 QY 481 IATAWCELQNLRELRLVWHEGIIKINPSATASATLGRVAAKMLGDVAASVSSCTAIDAESVT 540
 DB 487 IATAWCELQNLRELRLVWHEGIIKINPSATASATLGRVAAKMLGDVAASVSSCTAIDAESVT 546
 QY 541 QNSMRVITSTNTCYSRPLVLFYSYGENOGNIQOLGQENNELPTLEAVEPCSANHRRYFL 600
 DB 547 QNSMRVITSTNTCYSRPLVLFYSYGENOGNIQOLGQENNELPTLEAVEPCSANHRRYFL 606
 QY 601 GSGYALFENYFVKMVAADIIQIASTFVELMLTLEDEILPLSVYTKKEELRDVGVLGY 660
 DB 607 GSGYALFENYFVKMVAADIIQIASTFVELMLTLEDEILPLSVYTKKEELRDVGVLGY 666
 QY 661 EVARNOLHELK 672
 DB 667 EVARNOLHELK 678

RESULT 10

Q69408 ID Q69408 PRELIMINARY; PRT: 864 AA.
 AC Q69408; STRAIN=FC126;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GB homolog.
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=37108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC126;
 RX MEDLINE=94233711; PubMed=8178437;
 RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
 RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
 RT identification and expression by recombinant fowlpox viruses."
 RL Virology 200:484-493(1994).
 DR ENBL: U01887; AAA19447.1; --
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.

QY 604 YALFENTNFVMDADADIQIATSTFVELNLTLLREILPLSVYTKKEELRDVGVLDYAEVA 666
 DB 685 YVTFEDYAYVRKPLSEIELISAYVDLNTLLEDRBFLPLEVYVYTRAELEDTGLDYSIQ 744
 QY 664 RRNLQHELK 672
 DB 745 RRNLQHALK 753
 RESULT 15
 Q9OAP5 PRELIMINARY; PRT; 919 AA.
 ID Q9OAP5 PRELIMINARY; PRT; 919 AA.
 AC Q9OAP5; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glycoprotein B.
 OS Caprine herpesvirus 1 (goat herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae.
 RN NCBI_TaxID=39944;
 RN SEQUENCE FROM N.A.
 RC STRAIN=E/CH;
 RX MEDLINE=9921732; PubMed=10203465;
 RA Ros C., Belak S.;
 RT "Studies of genetic relationships between bovine, caprine, cervine,
 RT and rangifeline alphaherpesviruses and improved molecular methods for
 RT virus detection and identification";
 RL J. Clin. Microbiol. 37:1247-1253(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E/CH;
 RA Ros C., Belak S.;
 RT "Characterization of the glycoprotein B gene from ruminant
 RT alphaherpesviruses";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF078728; AAB46114.2; -;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 SQ SEQUENCE 919 AA; 100172 MW; A7DDCF03D13B4B66 CRC64;
 Query Match 50.0%; Score 1754.5; DB 12; Length 919;
 Best Local Similarity 52.0%; Pred. No. 1.2e-130;
 Matches 344; Conservative 92; Mismatches 204; Indels 21; Gaps 61
 QY 23 PSTQNTVS-----REVSVSSVQLSEESTFVLCPPPGVSTVIRLEPPKCPPEPRK 71
 DB 83 PSADGATAAAGNETAADDVRAVLRQAAGEHSQFFVCPPPSGATVYRLAPGRCPPEYL 142
 QY 72 ATEGEGIAILFKENISPYKFKVLTYYKNIQTWTGTYTTOITNRYTDRTVPVSEET 131
 DB 143 GRNTEGIGIYKENIAPYKFAHYKNIYVTTWSSGSIYAITNQHTDVRPVGSEIT 202
 QY 132 DLIDKGRCSKARYLRNNVYVEAFDRAGEKOVLLKPSKFNTPESRAWHTTTNETYVWG 191
 DB 203 DVVDKWRCLSKAEYLSGRKVVAFDRADPWEAPLKPSRLNAPGARGWHTTDEVHTVVG 262
 QY 192 SPWYRTGTSVNCIVEEMDARSVPYFYFAMANGDIANSFFYGLSPPEAAAPMGYPQD 251
 DB 263 SVGLYRTGTSVNCIVEEVEARSVPYDSFALSTGDIYMSPPFYGLR-EGAHREHTSYSPD 321
 QY 252 NFKOLDSYFSDMDLKRKASIPVKRNFLITSHETVGHWDWAPKTRVCSMTKWEVTEMLR 311
 DB 322 RFQOIEGYSRDLTSGRRAGDPVSRNFLTQHTVAVMDWPKRKNVCSLARAAEEMLR 381
 QY 312 ATVNGRYRFMARELSATFISNTTDFPNRIILGQICIKREAAEQIFRTKYNDSHVKVG 371
 DB 382 DESRGNRYFARALSATFVSDTHAFSLQHVPLSCVLEDAEAEEVRYRYNGTHVLG 441
 QY 372 HVQYFLAGGFTVAYQVPLSKSLAHMYLRELMDRNTDMDLVLNNKHATYKKNATSLSR 431

Db 442 KTDYTLARGGFVAFRPLLSNELAKLYLQELARSNRT---LDGVLGPR--LPAGAAPAGA 496
QY 432 LRRDIRNAPNRKITLDDTTAIKSTSSVOFAMLQFLYDHIQTHINDMFESRIATWCELQNR 491
Db 497 LRRARRAAPGGG---GGAGRVTVVSSAEFAALQFTYDHIQDHVNAME SRLATSWCLLQNK 553
QY 492 ELVLWHEGIIKINPSATNLGRVRAKMLGDVAVSSCTAIDAESVTLQNSMRVITSTN 551
Db 554 ERALWAEAKLNPSAAASALNRRRAARMILGDAMAVTYCRELGAGRVFIENSMR-MPGGA 612
QY 552 TCYSRPLVLFSYGENOCNIOGOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYN 611
Db 613 ACYSRPLVSPAYGNESEPLEGQGEDNELLRSRLVPCANHRRYFRFGADYVYENYA 672
QY 612 FVRWVDAADIQIASTFVELNLTLLEDREILPLSVYTKELRDVGVLDYAEVARRNOLHEL 671
Db 673 YVRVPLTETETISTFVDLNLTVLEDREFLPLEVYTRAE LADTGLLDYSEIQRRNOLHEL 732
QY 672 K 672
Db 733 K 733

Search completed: October 8, 2003, 17:07:07
Job time : 59.8445 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:35 ; Search time 12.1001 Seconds
(without alignments)
2611.703 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score: 3510

Sequence: 1 MHYFRNCIFFLIVLYLCTN.....DVGVLDAEYARRNQLHELK 672

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3510	100.0	865	1 VGLB_HSVMD	P18538 maret's dis
2	1740.5	49.6	980	1 VGLB_HSVEA	P18551 equine herp
3	1735.5	49.4	980	1 VGLB_HSVB	P28922 equine herp
4	1726	49.2	868	1 VGLB_HSV2	P09257 varicella-z
5	1720.5	49.0	979	1 VGLB_HSVEL	P25218 equine herp
6	1715.5	48.9	913	1 VGLB_HSVIF	P08355 pseudorabie
7	1707.5	48.6	919	1 VGLB_HSVE4	P17472 equine herp
8	1698	48.4	920	1 VGLB_HSVSM	Q04464 herpesvirus
9	1697.5	48.4	980	1 VGLB_HSVEL	P18050 equine herp
10	1693	48.2	904	1 VGLB_HSV2H	P08666 herpes slmp
11	1692	48.2	904	1 VGLB_HSV23	P06763 herpes slmp
12	1681	47.9	932	1 VGLB_HSVBC	P12640 bovine herp
13	1674	47.7	885	1 VGLB_HSV2S	P24994 herpes slmp
14	1666.5	47.5	917	1 VGLB_HSVB2	P12641 bovine herp
15	1663	47.4	903	1 VGLB_HSV1F	P06436 herpes slmp
16	1663	47.4	904	1 VGLB_HSV11	P10211 herpes slmp
17	1654	47.1	904	1 VGLB_HSV1P	P08665 herpes slmp
18	1651	47.0	904	1 VGLB_HSV1K	P06437 herpes slmp
19	1645	46.9	933	1 VGLB_HSV1L	Q04463 herpesvirus
20	1549.5	44.1	928	1 VGLB_HSVBP	P17471 bovine herp
21	1312	37.4	883	1 VGLB_ILTVS	P27415 infectious
22	1299.5	37.0	883	1 VGLB_ILTVT	P24904 infectious
23	1298.5	37.0	873	1 VGLB_ILTV6	Q02409 infectious
24	761	21.7	907	1 VGLB_HCMVT	P13201 human cytom
25	740	21.1	830	1 VGLB_HSV6U	P28864 human herpe
26	738.5	21.0	906	1 VGLB_HCMVA	P06473 human cytom
27	737	21.0	928	1 VGLB_HCMVS	P27171 murine cyto
28	736	21.0	908	1 VGLB_GPCMV	Q69024 guinea pig
29	732	20.9	830	1 VGLB_HSV6G	P36319 human herpe
30	726	20.7	830	1 VGLB_HSV6Z	P36320 human herpe
31	716.5	20.4	822	1 VGLB_HSV7J	P52352 human herpe
32	714.5	20.4	854	1 VGLB_HRCM6	P89053 rhesus cyto
33	694.5	19.8	857	1 VGLB_EBV	P03188 epstein-bar

RESULT 1

ID	VGLB_HSVMD	STANDARD;	PRT;	865 AA.
AC	P18538;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	GB.			
OS	Marek's disease herpesvirus (strain RB-1B) (MDHV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=33707;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89293086; PubMed=2544666;			
RA	Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;			
RT	"Nucleotide sequence and characterization of the Marek's disease			
RT	virus homologue of glycoprotein B of herpes simplex virus.";			
RL	J. Gen. Virol. 70:1789-1804 (1989).			
CC	- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; D13713; BAA02866.1;			
DR	InterPro; IPR000234; Glycoprot_B.			
DR	Pfam; PF00606; Glycoprotein_B; 1.			
DR	ProDom; PD000693; Glycoprot_B; 1.			
KW	Signal; Glycoprotein; Transmembrane.			
FT	SIGNAL 1 21 POTENTIAL.			
FT	CHAIN 22 865 GLYCOPROTEIN B.			
FT	DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 683 700 POTENTIAL.			
FT	TRANSMEM 709 729 POTENTIAL.			
FT	TRANSMEM 732 752 POTENTIAL.			
FT	DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).			
FT	CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 631 631 N-LINKED (GLCNAC. .) (POTENTIAL).			
SQ	SEQUENCE 865 AA; 98091 MW; B30E93C1AC5C5C63 CRC64;			

Query Match 100.0%; Score 3510; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 2.1e-249;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYLCTN.....DVGVLDAEYARRNQLHELK 672
|||||

Db 1 MHVFRNCIFFLVILYGTNSPSTQNTVSREVSVSVQSEESTYLCPVPVGVSTVIRL 60
 QY 61 EPPKCPKPKATWEGEGIAIFKFNISPYKFKVTLYYKNIQTTWTGTYRQITNRYT 120
 Db 61 EPPKCPKPKATWEGEGIAIFKFNISPYKFKVTLYYKNIQTTWTGTYRQITNRYT 120
 QY 121 DRTPVSEIETDLIDGKGRSSKARYLRNNVYAEFDRAGEQVLLKPSKFNTPESRAW 180
 Db 121 DRTPVSEIETDLIDGKGRSSKARYLRNNVYAEFDRAGEQVLLKPSKFNTPESRAW 180
 QY 181 HTTNEYTVWGSPIWYRTGTSVNCIVEEMDARSVPYFAMANGDIANTSPYGLSPPE 240
 Db 181 HTTNEYTVWGSPIWYRTGTSVNCIVEEMDARSVPYFAMANGDIANTSPYGLSPPE 240
 QY 241 AAAPMGYPQDNFKQLDSYFSDMLDKRRKASLPVKRNFLLTSFTVGMWAPKTRVCSM 300
 Db 241 AAAPMGYPQDNFKQLDSYFSDMLDKRRKASLPVKRNFLLTSFTVGMWAPKTRVCSM 300
 QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQCICKREAAISQIFR 360
 Db 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQCICKREAAISQIFR 360
 QY 361 TKYNDSHVKGVOYFLALGCFIVAYQPVLSKSLAHMYLRELMDNRDNTDMLDLVNNKHA 420
 Db 361 TKYNDSHVKGVOYFLALGCFIVAYQPVLSKSLAHMYLRELMDNRDNTDMLDLVNNKHA 420
 QY 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHTQTHINDMFSR 480
 Db 421 IYKKNATSLSLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHTQTHINDMFSR 480
 QY 481 IATAWCEQLQRELVLHHEGIIKIPASATASATLGRVAAKMLGDVAVSSCTAIDAESVTL 540
 Db 481 IATAWCEQLQRELVLHHEGIIKIPASATASATLGRVAAKMLGDVAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLFSGENQIGQVGENNLEPTEAVEPCSANHRRYFLF 600
 Db 541 QNSMRVITSTNTCYSRPLVLFSGENQIGQVGENNLEPTEAVEPCSANHRRYFLF 600
 QY 601 GSGVALFENFVKWDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDDYA 660
 Db 601 GSGVALFENFVKWDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDDYA 660
 QY 661 EVARNQOLHELK 672
 Db 661 EVARNQOLHELK 672

RESULT 2
 VGLB_HSVEA STANDARD: PRT; 980 AA.
 ID VGLB_HSVEA
 AC P18551;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33.
 OS Equine herpesvirus type 1 (strain ABL) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae.
 OX NCBI_TaxID=10328;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bonass W.A., Elton D.M., Stocks J.M., Killington R.A.,
 RA Meredith D.M., Halliburton I.W.;
 RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

or send an email to license@isb-sib.ch.
 CC EMBL; M36298; AAA46068.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86 POTENTIAL.
 FT CHAIN 87 980 GLYCOPROTEIN B.
 FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 853 870 POTENTIAL.
 FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 980 AA; 109736 MW; 264D273CED70B5A1 CRC64;
 Query Match 49.6%; Score 1740.5; DB 1; Length 980;
 Best Local Similarity 48.1%; Pred. No. 1.3e-119;
 Matches 343; Conservative 113; Mismatches 188; Indels 69; Gaps 11;
 QY 19 TNSPSTQNTVS-----REVSVSVQSEESTYLCPVPVGVST 56
 Db 88 TTPSPPTSTPSTASHGTVDPFTLLPTPTDPLRLAVRESGILAEQDFYTCPPPTGT 147
 QY 57 VIRLEPPKCPKPKATWEGEGIAIFKFNISPYKFKVTLYYKNIQTTWTGTYRQIT 116
 Db 148 VVRIEPPKCPKDLGNETFEGIAIFKFNISPYKFKVTLYYKNIQTTWTGTYRQIT 207
 QY 117 NRYDTRTPVSEIETDLIDGKGRSSKARYLRNNVYAEFDRAGEQVLLKPSKFNTP 176
 Db 208 DRYNDRVPSVSEIEFLGLIDSKGKSKAEYLRNIMHAYHDEDEVELDLVPSKATPG 267
 QY 177 SRAWHTTNET--YTVMGSPWIYRTGTSVNCIVEEMDARSVPYFAMANGDIANTSPFY 234
 Db 268 ARAWQTTNDTTSYVGW-MPWRHYTSTSVNCIVEEARSVPYDVSFALSTGDIYASPFY 326
 QY 235 GLSPPEAAA--EPMGYPQDNFKQLDSYFSDMLDKRRKASLPVKRNFLLTSFTVGMW 292
 Db 327 GL---RAAARIEHNSYAQDSFRQVEGYRPRDLSKQAEPEFTKNFTTTPHVTVSWNTE 383
 QY 293 KTTKRVCSMTKWKVEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQCICKRE 352
 Db 384 KKEVEACTLTKWKEVDELVRDEFGRGTYFTIRSSISFTISNTQPKLESAPLTCVSK 443
 QY 353 AAIEQIFRTKYNDSHVKGVOYFLALGCFIVAYQPVLSKSLAHMYLRELMDNRDNTD 412
 Db 444 EAIDSIVKKQYESTHVPFSGDVEYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 503
 QY 413 DLYNNKAIYKKNATSLSLRRDIRNAPNRKITL-----DTA----- 451
 Db 504 NLLNPN-----ANNNNNTTTRRRRLSVPEPTQDGVHREQLHRLHRLHRAVEATAGT 559
 QY 452 -----IKTSSVQFAMQLFLYDHTQTHINDMFSRIATACWELNRELVLHHEGII 502
 Db 560 NVYAKQLELTKTSSIEFAMQLQFAYDHIQSHVNEMLSRISATACWELNRELVLHHEGII 619
 QY 503 NPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTQLQNSMRVITSTNTCYSRPLV 562
 Db 620 NPSAIVSATLDERVAARVLGDVIAITHCAKIEG-NVYLQNSMRSMDS-NTCYSRPPT 677
 QY 563 Y---GENQIGQVGENNLEPTEAVEPCSANHRRYFLGSGYALFENFVKWDA 619
 Db 678 ITKANNRGSGEQVGENEIEFTKRLIEPCALNKKRYFKGKYEIVYENTFVYKVPPT 737
 QY 620 DIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDDYA EAVARRNOLHELK 672

738 ELIEVSTVVELNLTLEDEFLPEVYTRAELEDGLDYSEIQRRNQLHALR 790

Db 88 TTSPSTPTSTSTSHSGTVDPTLLPTTDPDLRAVRESGILAEADGDFYTCPPPTGST 147

QY 57 VIRLEPPKPEPRKATENGEGIAILFKENISPIYKFKVLYYKNIQIOTTTWGTYYRQIT 116

Db 148 VVRIEPPRTCPDFDLGRNFTEGIAVFKENIAPYKFRANVYKDIIVTVRWKGYSHTSLS 207

QY 117 NRYTDRTPSVIEIEITLDIDKGRCSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPPE 176

Db 208 DRYNDVPSVVEIEFLIDSKGCSKAEYLRDNIMHAYHDEDEVELDLVPSFAFPG 267

QY 177 SRAWHTNET--YTVWGSPWIYRTGYSVNCIVEEMDARSFVYSYFAMANGDIANISPPY 234

Db 268 ARAWQTTNDTTSYVGW-MPWRHYTSTVNCIVEEARSVYDPSFALSTGDIIVASPPY 326

QY 235 GLSPPEAAA--EPMGYPODNFKOLDYSFSDLDKREKASLPVKRNFILTSHTFTVGNDRAP 292

Db 327 GL---RAAARIEHNYSQAERFQVGRPRDLDKLAEEPTVKFNFTTTPHYTVSNMTE 383

QY 293 KTRVCSMTKWKVEEMLRATVNGRYRFMARLSATFISNTTFFDPNRIILQCCKREAE 352

Db 384 KKVEACTLTKWEYDELVRDEFRGSRFYRISISSTFISNTTQFKLESAPLTCVSKEAK 443

QY 353 AAIEQIFRTKYNDSHVKGVOYFLALGQFIVAYQPVLSKLAHMYLRELMRDNRTDEML 412

Db 444 EADISYVKKQYESTHVFGSDVEYVLARGFLIAFRPMLSNELARLYLNLVSNRTYDLK 503

QY 413 DLYNNKHAITYKKNATSLRLRDIRNAPNRKTYLD-----DTTA----- 451

Db 504 NLLNPN-----ANNNTTTRRRSLLSVPPEQPTQGVHREQILHLRHKRAVATAGTSS 559

QY 452 -----IKTSSVQFAMLOFLYDHIQTHINDMFSRIATAWCELQNLRELVLWHEGKI 502

Db 560 NYTAKOLELKITTSIEFAMLOFAYDHIQSHVNEMLSRITATWCIQNKERTLWNEWVKI 619

QY 503 NPSATASATLGRVAAKMLGDDVAVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLVLF 562

Db 620 NPSAIVSATLDERVAARVLGDVIAIFHCAKIEG-NVYLQNSMRSDS-NTCYSRPEVPT 677

QY 563 Y---GENOGNIOGLOKNNELLPTLEAVEPCSNHRRYVFLFGSGYALFENYFNKVDMA 619

Db 678 ITKNANRRSIGQLGEENEIFETFKLIEPCALNQKRYFKFQKVEYVYVTFYRKPVPPT 737

QY 620 DIQIATSFVELNLTLEDEILPLSVYTKEEELRDLGVLDYAEVARRNQLHELK 672

Db 738 EIEVISTVVELNLTLEDEFLPEVYTRAELEDGLDYSEIQRRNQLHALR 790

RESULT 4

VLGB_VZVD STANDARD; PRT; 868 AA.

AC P09257; 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycoprotein B precursor (Glycoprotein II).

GN 31.

OS Varicella-zoster virus (strain Dumas) (VZV).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=10338; [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=86306657; PubMed=3018124;

RX Davidson A.J., Scott J.E.;

RA "The complete DNA sequence of varicella-zoster virus.";

RT J. Gen. Virol. 67:1759-1816(1986).

RL -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC

Db 88 TTSPSTPTSTSTSHSGTVDPTLLPTTDPDLRAVRESGILAEADGDFYTCPPPTGST 147

QY 57 VIRLEPPKPEPRKATENGEGIAILFKENISPIYKFKVLYYKNIQIOTTTWGTYYRQIT 116

Db 148 VVRIEPPRTCPDFDLGRNFTEGIAVFKENIAPYKFRANVYKDIIVTVRWKGYSHTSLS 207

QY 117 NRYTDRTPSVIEIEITLDIDKGRCSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPPE 176

Db 208 DRYNDVPSVVEIEFLIDSKGCSKAEYLRDNIMHAYHDEDEVELDLVPSFAFPG 267

QY 177 SRAWHTNET--YTVWGSPWIYRTGYSVNCIVEEMDARSFVYSYFAMANGDIANISPPY 234

Db 268 ARAWQTTNDTTSYVGW-MPWRHYTSTVNCIVEEARSVYDPSFALSTGDIIVASPPY 326

QY 235 GLSPPEAAA--EPMGYPODNFKOLDYSFSDLDKREKASLPVKRNFILTSHTFTVGNDRAP 292

Db 327 GL---RAAARIEHNYSQAERFQVGRPRDLDKLAEEPTVKFNFTTTPHYTVSNMTE 383

QY 293 KTRVCSMTKWKVEEMLRATVNGRYRFMARLSATFISNTTFFDPNRIILQCCKREAE 352

Db 384 KKVEACTLTKWEYDELVRDEFRGSRFYRISISSTFISNTTQFKLESAPLTCVSKEAK 443

QY 353 AAIEQIFRTKYNDSHVKGVOYFLALGQFIVAYQPVLSKLAHMYLRELMRDNRTDEML 412

Db 444 EADISYVKKQYESTHVFGSDVEYVLARGFLIAFRPMLSNELARLYLNLVSNRTYDLK 503

QY 413 DLYNNKHAITYKKNATSLRLRDIRNAPNRKTYLD-----DTTA----- 451

Db 504 NLLNPN-----ANNNTTTRRRSLLSVPPEQPTQGVHREQILHLRHKRAVATAGTSS 559

QY 452 -----IKTSSVQFAMLOFLYDHIQTHINDMFSRIATAWCELQNLRELVLWHEGKI 502

Db 560 NYTAKOLELKITTSIEFAMLOFAYDHIQSHVNEMLSRITATWCIQNKERTLWNEWVKI 619

QY 503 NPSATASATLGRVAAKMLGDDVAVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLVLF 562

Db 620 NPSAIVSATLDERVAARVLGDVIAIFHCAKIEG-NVYLQNSMRSDS-NTCYSRPEVPT 677

QY 563 Y---GENOGNIOGLOKNNELLPTLEAVEPCSNHRRYVFLFGSGYALFENYFNKVDMA 619

Db 678 ITKNANRRSIGQLGEENEIFETFKLIEPCALNQKRYFKFQKVEYVYVTFYRKPVPPT 737

QY 620 DIQIATSFVELNLTLEDEILPLSVYTKEEELRDLGVLDYAEVARRNQLHELK 672

Db 738 EIEVISTVVELNLTLEDEFLPEVYTRAELEDGLDYSEIQRRNQLHALR 790

RESULT 4

VLGB_VZVD STANDARD; PRT; 868 AA.

AC P09257; 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycoprotein B precursor (Glycoprotein II).

GN 31.

OS Varicella-zoster virus (strain Dumas) (VZV).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=10338; [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=86306657; PubMed=3018124;

RX Davidson A.J., Scott J.E.;

RA "The complete DNA sequence of varicella-zoster virus.";

RT J. Gen. Virol. 67:1759-1816(1986).

RL -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

Db 88 TTSPSTPTSTSTSHSGTVDPTLLPTTDPDLRAVRESGILAEADGDFYTCPPPTGST 147

QY 57 VIRLEPPKPEPRKATENGEGIAILFKENISPIYKFKVLYYKNIQIOTTTWGTYYRQIT 116

Db 148 VVRIEPPRTCPDFDLGRNFTEGIAVFKENIAPYKFRANVYKDIIVTVRWKGYSHTSLS 207

QY 117 NRYTDRTPSVIEIEITLDIDKGRCSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPPE 176

Db 208 DRYNDVPSVVEIEFLIDSKGCSKAEYLRDNIMHAYHDEDEVELDLVPSFAFPG 267

QY 177 SRAWHTNET--YTVWGSPWIYRTGYSVNCIVEEMDARSFVYSYFAMANGDIANISPPY 234

Db 268 ARAWQTTNDTTSYVGW-MPWRHYTSTVNCIVEEARSVYDPSFALSTGDIIVASPPY 326

QY 235 GLSPPEAAA--EPMGYPODNFKOLDYSFSDLDKREKASLPVKRNFILTSHTFTVGNDRAP 292

Db 327 GL---RAAARIEHNYSQAERFQVGRPRDLDKLAEEPTVKFNFTTTPHYTVSNMTE 383

QY 293 KTRVCSMTKWKVEEMLRATVNGRYRFMARLSATFISNTTFFDPNRIILQCCKREAE 352

Db 384 KKVEACTLTKWEYDELVRDEFRGSRFYRISISSTFISNTTQFKLESAPLTCVSKEAK 443

QY 353 AAIEQIFRTKYNDSHVKGVOYFLALGQFIVAYQPVLSKLAHMYLRELMRDNRTDEML 412

Db 444 EADISYVKKQYESTHVFGSDVEYVLARGFLIAFRPMLSNELARLYLNLVSNRTYDLK 503

QY 413 DLYNNKHAITYKKNATSLRLRDIRNAPNRKTYLD-----DTTA----- 451

Db 504 NLLNPN-----ANNNTTTRRRSLLSVPPEQPTQGVHREQILHLRHKRAVATAGTSS 559

QY 452 -----IKTSSVQFAMLOFLYDHIQTHINDMFSRIATAWCELQNLRELVLWHEGKI 502

Db 560 NYTAKOLELKITTSIEFAMLOFAYDHIQSHVNEMLSRITATWCIQNKERTLWNEWVKI 619

QY 503 NPSATASATLGRVAAKMLGDDVAVSSCTAIDAESVTLQNSMRVIT

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; X04370; CAA27914.1; -.
DR PIR; E27214; VGBE31.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 868 GLYCOPROTEIN B.
FT DOMAIN ? 688 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 680 695 POTENTIAL.
FT TRANSMEM 701 720 POTENTIAL.
FT TRANSMEM 724 744 POTENTIAL.
FT DOMAIN 745 868 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 557 557 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 868 AA; 98066 MW; 58B8E1346F902 CRC64;

Query Match 49.2%; Score 1726; DB 1; Length 868;
Best Local Similarity 50.5%; Pred. No. 1.3e-118;
Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;

QY 31 REVSSVOLSEESFYLCPPVGVSTVIRLPPKCPKPRKATEGEGIAILFKENISPY 90
DB 41 REAIKSDQAEKPTFYCPPTGTTIVRLPEPTRCPDHLGKNTGEGIAVVKENIAY 100

QY 91 KFKVLYXKNIIQTWTGTYRQITNRYTRDTPVSEIEITDLIDGKRCSSKARYLRNN 150
DB 101 KFKATVYKDVIVSTAWAGSSVQITNRYADRPVPIVSEITDIDKFKCCKSKATYVRN 160

QY 151 VYVEAFDRDAKEQVLLKPSKENTPESRAWHTTNTYVWGSPIYRTGTSVNCIVEMD 210
DB 161 HKVEAFNEKPNQDMPLIASYNSVGSKAHTTNDTYVAGTPTGYRTGTSVNCIIEVE 220

QY 211 ARSVFPYVFAMANGDIANISPFYGLSPPEAAAPMGYPQDNFKOLDSYFMDLKRKA 270
DB 221 ARSIFPYSFGLSTGDIYIMSPFFGLR-DGAYREHSNTAMDFHGFQYRQDLTRALL 279

QY 271 SLVPKRNFLITSHFTVGMWNAKPTRVCSMTKWKVEMLRATVNGRVKEMARELSATFI 330
DB 280 E-PAARNEFLVPLHTVGNWPKPKRTEVCSLVKREVEDVVRDEYAHNFRFTMKTLSFTFI 338

QY 331 SNTTFEDPNRIILGOCIKREAAEQIFRTKYNDSHVKVGHVQVFLALGGFIVAYQVPL 390
DB 339 SETNEFNQIHLSCVKEARAIINRYTTRYNSSHVGTGDIQTYLARGGFVVFQPLL 398

QY 391 SKSLAHMYLRELNRNRTDEMLDLVNNKHAYIKKNATSLSLRLRDINAPNKRKITLDDTT 450
DB 399 SNSLARLYQLVRE-----NTNHS-PQKHPTNRSRSRV-----PVELRANR 441

QY 451 AIKTSSTVOFAMQLFYDHIQTHINDMPSRIATANCELQNRRLVWHEGIKINPATSASA 510
DB 442 TITTTSSVEFAMQLQTYDHIQTHVNEMLARISSSQCLQNRRLARLWGLFPINPASALAT 501

QY 511 TIGRVAAKMLGDVAVSCITAIDAES-VTLQNSMRVITSTWTCYSRPLVLFSGENQGN 569
DB 502 ILDRVKARILGDVTSVNCNCPGLSDTRIILQNSMRVSGSTRCSYRPLISIVLSNGSGT 561

QY 570 IQQGLGENNELPTLEAVPCSANRRYFLFGSGYALFENYFVKWVDAADAOIASTFVE 629
DB 562 VESQGLGTNELLMSRDLLEPCVANKRFLFGHHVYVYEDYRYVREIAVHDVGMISTYVD 621

QY 630 LNLTLLEDREILPLSVYTKKELRDVGVLDAEYAVARNQLHELK 672
DB 622 LNLTLKOREFPLQVYTRDELDRDTGLLDYSIQRRNQMSLR 664
```

RESULT 5

```
VGLB_HSVEL STANDARD; PRT; 979 AA.
AC P25218;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR gp14 OR 33.
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10330;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236317; PubMed=1692002;
RA Guo P.;
RT "Characterization of the gene and an antigenic determinant of equine
RT herpesvirus type-1 glycoprotein 14 with homology to gb-equivalent
RT glycoproteins of other herpesviruses.";
RT Gene 87:249-255(1990).
RL Gene 87:249-255(1990).
CC 1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34861; AAA46086.1; -.
DR PIR; JH0109; JH0109.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 979 GLYCOPROTEIN B.
FT DOMAIN 87 851 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 852 869 POTENTIAL.
FT DOMAIN 870 979 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 979 AA; 110331 MW; 9A19866B791C5B36 CRC64;

Query Match 49.0%; Score 1720.5; DB 1; Length 979;
Best Local Similarity 47.7%; Pred. No. 3.9e-118;
Matches 340; Conservative 113; Mismatches 191; Indels 69; Gaps 11;

QY 19 TNSPSTQNTVS-----REVSSVOLSEESFYLCPPPVGST 56
DB 88 TTPSPSTPTSMSTSHSGTVDPTLLPTPTDPLRLAVRESGILAECDGDFYCPPTGST 147

QY 57 VIRLEPPPKRCPKPRKATEGEGIAILFKENISPYKFKVLYXKNIIQTWTGTYRQIT 116
DB 148 VYRIEPPRTCPKFLGRNFTGEGIAVFKENIAPYKFRANVYKDIIVTVRWKGYSHSLS 207

QY 117 NRYTDRTPVSEIEITDLIDGKRCSSKARYLRNNVVEAFDRDAKEQVLLKPSKENT 176
DB 208 DRYNDRVPVSEIEIFGLIDSKCKSKAEYLRNIMHHAYHDDHDEVDLVPKFPATPG 267

QY 177 SRAWHTTNET--YTVWGSPIYRTGTSVNCIIEEMDARSVFPYSFAMANGDIANISPY 234
```

Db 268 ARAQWNTDTSYVGW-MPRHRTSTSVNCIVEEARSVPYDPSFALSTGDIIVYASPFY 326
 Qy 235 GLSPPEAAA--EPNGYPQDNFKOLDYSFMDLKKRRKASLPVKNFLITSHFTVGNWDAP 292
 Db 327 GL---RAAARIEHNSYAQERFQVGYRPRDLKSLQAEPEVTKNFITTHVTVSNWTE 383
 Qy 293 KTRVCSKWKKEVTEMLRATVNGRYRFBARELSATISNTTDPNRIILGQCIRKAE 352
 Db 384 KKEACFLTAKWEVDELRVDFRGSYRFTIRSISSYFISNTQFKLESAPLTCVSKAK 443
 Qy 353 AAIEQIPRTYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYRLMNRNRTDEML 412
 Db 444 EADSIVKKQYESTHVSVDGYEYVLAGRGFLIAPRPMLSNELARLYLNELVRSNRTYDLK 503
 Qy 413 DLVNNKHAIYKKNATSLRSLRDRIRNAPNKKIULD-----DTTA----- 451
 Db 504 NLLPNP-----ANNNTTTRRRSSLLSVPEPQTDGQVHRQILHRLHRAVEATAGDSS 559
 Qy 452 -----IKSTSSVQFAMQLFDYDHIOTHINDMESRIATAMCELQNRRLVLMHEGKI 502
 Db 560 NVTAKOLELKTSSIEFAMQLFAYDHIQSHVNEMLSRATAMCPLQNKERPLNEMVKI 619
 Qy 503 NPSATASATLGRVAAKMLGDVAASVTAIDAESVTLQNSMRVITSTNTCYSRPLVLS 562
 Db 620 TPSAIVSATLDERVAARVLGDVIAITHCAKIEG-NVYLQNSMRSDS-NTCYSRPPVTF 677
 Qy 563 Y---GENQNIQQLGNNELLPTLEAVEPCSANHRRYFLGSGYALFENYFVNMVDA 619
 Db 678 ITRNANRSGTEGOLGSENEIFTERKLTIEPCALNQKRYFKGKYVYVYNYTFVRKVPPT 737
 Qy 620 DIQIASTFVELNLTLLDRETLPLSVYTKBELRDVGVDYAEVARRNOLHELK 672
 Db 738 EIEVISTYVELNLTLLDREFLPLEVYTRAELEDTGLLDYSEIQRRNOLHALR 790

RESULT 6

VGLB_PRIV VGLB_PRIV STANDARD; PRY; 913 AA.
 AC P08355;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein GII precursor.
 OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 ON NCBI_TaxID=31523;
 RX MEDLINE-87284141; PubMed-3039163;
 RA Robbins A.K., Dorney D.J., Wathen M.W., Whealy M.E., Gold C.,
 RA Watson R.J., Holland L.E., Weed S.D., Levine M., Glorioso J.C.,
 RA Enquist L.W.;
 RT "The pseudorabies virus gII gene is closely related to the gB
 glycoprotein gene of herpes simplex virus.";
 RL J. Virol. 61:2691-2701(1987).
 RN [2]
 RP SEQUENCE OF 847-913 FROM N.A.
 RX MEDLINE-89279296; PubMed-2543777;
 RA Simon A., Mettenleiter T.C., Rziha H.J.;
 RT "Pseudorabies virus displays variable numbers of a repeat unit
 adjacent to the 3' end of the glycoprotein gII gene.";
 RL J. Gen. Virol. 70:1239-1246(1989).
 RN [3]
 RP EXPORT PATHWAY.
 RX MEDLINE-90219190; PubMed-2157862;
 RA Whealy M.E., Robbins A.K., Enquist L.W.;
 RT "The export pathway of the pseudorabies virus gB homolog gII involves
 oligomer formation in the endoplasmic reticulum and protease
 processing in the Golgi apparatus.";
 RL J. Virol. 64:1946-1955(1990).
 CC -I- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M17321; AAA47465.1; -;
 DR EMBL; D00464; BAA00359.1; -;
 DR PIR; A29159; VGBEPS.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 40
 FT CHAIN 41 913 GLYCOPROTEIN GII.
 FT DOMAIN 41 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 819 POTENTIAL.
 FT DOMAIN 820 913 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 516 516 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 913 AA; 100234 MW; 5D56D0235E856437 CRC64;
 Query Match 48.9%; Score 1715.5; DB 1; Length 913;
 Best Local Similarity 50.7%; Pred. No. 8.1e-118;
 Matches 340; Conservative 105; Mismatches 184; Indels 41; Gaps 12;
 Qy 18 GTNSPSPTQ--NVTSEVWSVOLSEESTFYLCPPVGVSTVIRLEPKCPKPEKATEW 75
 Db 95 GPSEAPDGEYGLDARTAVRAA--ATERDRFYVCPSPSGSVVLEPQACPEYSQGRNF 152
 Qy 76 GEGIAILFKENISPKFKVLYKNIOTTTWTGTYRQINRYTRDTPVSEIEITDLD 135
 Db 153 TEGIAVLFKENIAPHKFAHYKNIIVTTVWSGSTYAAITNRTDRVPVQVEITDVID 212
 Qy 136 GKGRCSSKARYLRNNVVEAFDRDAGEKQVLLKSKFWTPSPESRAWHHTTNEYTVWGSWI 195
 Db 213 RRGKCVSKAEYVRNNHKVTAFDRENPEVDLRPSRLNALGTGRHHTTNDYTKIGAAGF 272
 Qy 196 YRTGTSVNCIVEEMDARSVPFYSFAMANGDANISPEYGLSPPEAAAPMGYPQDNFKQ 255
 Db 273 YHTGTSVNCIVEEARSVPYDPSFALSTGDIIVYMSFYGLR-EGAHGEHIGYAPGRFQ 331
 Qy 256 LDSYFMDLDRKRRKASLPVKNFLITSHFTVGNWDAPKTRVCSKWKKEVTEMLR-ATV 314
 Db 332 VEHYPIIDLDSRLRAESVTNRLTPHTFTVWDWAPKTRRVCSLAKWREAEEMTRDETR 391
 Qy 315 NGRYRFARELSATISNTTDPNRIILGQCIRKAEAEAEQIFRTYNDSHVKG-VH 373
 Db 392 DGSFRTSRALGASFVSDVTQDLQVRHLGDCVLEASAEADAIYRRYNTSHVLAGDRP 451
 Qy 374 QYFLALGGFIVAYQPVLSKSLAHMYRLMNRNRTDEMLDVLNNKHAIYKKNATSLRSLR 433
 Db 452 EYVLARGGFVAFRPLISNELAQLYAREL-----ERLGLAG---VVGPAAPAAARRAR 501
 Qy 434 RD-----IRNAPNRKITLDDTTAIAKSTSSVQFAMQLFDYDHIOTHINDMFSRIA 482
 Db 502 RSPGPAGTPEPPAVNGTGLHRLT-----TGSAEFARLQFTYDHIQAHVNDMLGRIA 552
 Qy 483 TAWCELQNRRLVLMHEGKIINPSATASATLGRVAAKMLGDVAASVCSCTAIDASVTIQN 542
 Db 553 AAWCELQNKDRTLWSENRLNPSAVATAALQGRVSARMGLGDVAISRCVEVRG-GVYVQN 611
 Qy 543 SMRYITSTNTCYSRPLVLSYGENQNIQQLGNNELLPTLEAVEPCSANHRRYFLGSG 602
 Db 612 SMRYPGERGTCYSRPLVTTFE-H-NGTVIEGQLGDDNELLISRDLEPCTGNHRRYFKLGS 670

QY 603 GYALFENYFVKVDAADIAQIASTFVFLNLTLLDREILPLSVYTKRELDRGVLDYAEV 662
DB 671 GYVIEDYNYRMVYEP--ETISTRVTLNLTLLDREILPLSVYTKRELDRGVLDYAEV 728
QY 663 ARRNQLHEK 672
DB 729 QRRNLHAK 738

RESULT 7
VGLB_HSV4
ID VGLB_HSV4 STANDARD; PRT; 919 AA.
AC P17472;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Glycoprotein B precursor.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Alphaherpesvirinae; Varicelloviruses;
RN NCBI_TaxID=10333;
RP SEQUENCE FROM N.A.
RX MEDLINE=89125704; PubMed=2915378;
RA Rigbio M.P., Cullinane A.A., Onions D.E.;
RT Identification and nucleotide sequence of the glycoprotein gb gene of equine herpesvirus 4.;
RL J. Virol. 63:1123-1133(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation. use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; M26171; AAA46106.1; ALT_INIT.
DR PIR; A31880; VGBEQH.
DR InterPro: IPR00234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 919 GLYCOPROTEIN B.
FT DOMAIN 41 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 809 POTENTIAL.
FT DOMAIN 810 919 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 919 AA; 103710 MW; C77E48F26C37BC7B CRC64;

Query Match 48.6%; Score 1707.5; DB 1; Length 919;
Best Local Similarity 47.9%; Pred. No. 3.2e-117;
Matches 336; Conservative 117; Mismatches 195; Indels 53; Gaps 11;
QY 20 NSSPST-----QNTSREVSSVQLSEES-----TFYLCPPPVGVSTVIRLE 61
DB 34 SSQPTPASTQSAKTVDQTLPTPTDPLRLAVRESSILAEDGDFTCPPPGSTVIRIE 93
QY 62 PPRKCPKPRKATGEWGEGIALFKENISPKFKVLYKKNIQTWTGTTRQITNRYND 121

DB 94 PPSCPKFDLGNFTGIAVIFKRNATYKIDVTVTKVWKGYSHSLSDRYND 153
QY 122 RPTVSTEEITDIDKGRGSSKARYLNRYVVEAFDRDAGEKQVLLKPSKFTWTPSRAWH 181
DB 154 RVPVSVEEITLDSKKGSSKAEYLRONIMHAHDDDEVELDLVPSKFATPGARAWQ 213
QY 182 TTNET--YTVWGSPWYRTGTVCNIVEMDARSVPYSYFAMANGDANISPPFVGLSP 239
DB 214 TTNDTTSVGM-MPWRHYTSTSVNCIVEEARSVPYDPSFALSTGDIVYTSPPFVGLR-S 271
QY 240 EAAEPMGYPQDNFKQLDSYFMDLDRKASLPVKRNLITSHPTVGMWNAKTRVCS 299
DB 272 AAQLEHNSYAQERFQEGYQPRDLDSKLQAGEPVTNFTTTPHTVTSWNTKIEACT 331
QY 300 MTKKKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDNRILGQCICKREAAAIQIF 359
DB 332 LTKWKEVDLVRDEFGRGYSRTIRSISSFTSNTQKLEADAPLTCVSKAKDAISY 391
QY 360 RTKYNDSHVKGVOYELAGGFVAYQVLSLAHMYLRELNRDNRDTEMLDVLNKH 419
DB 392 RKQYESTHVSGDVEFYLAGRGFLAFRPMISNELARLYLNELVRSNRTYDLKNLL-NPN 450
QY 420 AIYKKNATS-----LSRLRDIRNAPN-----RKITLDDTTAKS 454
DB 451 ANHNTNRRSLLSIPEPTQBSLHREQLHLHRAVEAANSTNSSNVTAKQLELIK 510
QY 455 TSSVQFAMQLFDYDHTQTHINDMFSRIATAWCELQRELVLHHEGKINPSATASATLGR 514
DB 511 TSSIEFAMQFADYDHTQSHVNEMLSRATATWCTLQNKERTLWNEVKVAPSALVATLDE 570
QY 515 RVAAKMLGDVAVSSCTAIDAESVTLQNSRVITSTNTCYSRPLVLFSGYEN---QGNIQ 571
DB 571 RVAARVLGDVIAITHCVKIEG-NVYLQNSMR-SSDSNTCYSRPPVTFITTKNANSRGTE 628
QY 572 GOLGENNELLPLEAVEPCSAHRRYFLFGSGYALFENYFVKVMDAADIQIASTFVELN 631
DB 629 GOLGEENEVYTERKLTPECAINQKRYFKGKVEYVYENYTKRVPPTIEVISTVYELN 688
QY 632 LTLLDREILPLSVYTKRELDRGVLDYAEVARNQLHEK 672
DB 689 LTLLDREILPLSVYTKRELDRGVLDYAEVARNQLHEK 729

RESULT 8
VGLB_HSVSM
ID VGLB_HSVSM STANDARD; PRT; 920 AA.
AC Q04464;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Glycoprotein B precursor.
OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Alphaherpesvirinae.
RN NCBI_TaxID=10353;
RP SEQUENCE FROM N.A.
RX MEDLINE=93228440; PubMed=8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gb gene homologs of two platyrrhine monkey alpha-herpesviruses.";
RL Arch. Virol. 129:167-182(1993).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DB EMBL: M95786; AAA43841.1; -
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28: POTENTIAL.
FT CHAIN 29 920 GLYCOPROTEIN B.
FT DOMAIN 29 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 755 POTENTIAL.
FT TRANSMEM 761 781 POTENTIAL.
FT TRANSMEM 784 804 POTENTIAL.
FT DOMAIN 805 920 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;

Query Match 48.4%; Score 1698; DB 1; Length 920;
Best Local Similarity 50.2%; Pred. No. 1.6e-116;
Matches 322; Conservative 93; Mismatches 201; Indels 26; Gaps 3;

QY 31 REVSSVOLSEESTFYLCPPPPGVSVIRLEPPKPCPEPRKATWEGGAILFKENISPY 90
DB 109 RESVQYRAENATSMFYVPPPTGATVVOQPEPRPCPDVAAGNFTEGIAVIFKENIAPY 168
QY 91 KEKVTLYKNIIOITTTGGTYTQITNRYTDRPVSTEEITDIDGKRCSSKARYLRNN 150
DB 169 KFTATMYKEITVTQWQSGRYQLTGLYNDRAPVPEEITDINAKGLCRSDVTVRSQ 228
QY 151 VYVEAFDRDAGEKQVLKPKSFNTPESAWHTTNTYTVGSPWIVTGTSVNCLVEEMD 210
DB 229 RRVYAYDRDWEGRVGLVPKSTSTPSNRSGWYTTDRMYPNAHAGFYKAGTIVNCIVEE 288
QY 211 ARSVFYSYFAMANGDANTISPYGLSPPEAAEPKGYPODNFKQDYSFMDLKRKA 270
DB 289 ARSAYPSYFNLATGDFVYVSPYGLG-EDAHREYNAYSADRFKQDGFPPRLDSEGA 347
QY 271 SLPKRNLITTSHTVGMNAPKTRVCSMTKKEVTEMLRATVNGRYRPMARELSATFI 330
DB 348 PEPVVRMLTTPFTIGDWKPKDPSPVCSVKQVEEMRAEYSGSTFRSTSSLSATFT 407
QY 331 SNTTEFDPNRIILQCCIKREAEAAIEQIFRTKYNDSHVKYGVHVOYFALAGGFIVAYQPV 390
DB 408 TNYTQYPPQRIELSDCAREAAQAVDAIYARRYNASHVKYVGLQYLAQGFELVYQPLI 467
QY 391 SKSLAHMYRLMDNRTDEMDLVNKNHAIYKKNATSLSLRDRIRNAPRKILTDIT 450
DB 468 SNSLAHYLRE-----AARALEPAPLPTTTPAPE---AAGSRG 502
QY 451 AIKSTSSVQFAMQLQFLYDHIOTINDMFSRIATAWCELOQRELVLWHEGIKINPSATASA 510
DB 503 TLSTQTSQVEFARQFTYDHLQKHVNEMLGRIAAWQLOQELVLWNEARKLPNATASA 562
QY 511 TLGRVAAKMLGDVAAVSSCTADAESVTTLQNSRVITSTNTCYSRPLVLFSGYGENOGNI 570
DB 563 TVGRVGMMLGDVMAVSTCIAPVPHVIMQNSMRLPARKPKTCYARPLVSFRVDEGELI 622
QY 571 QGQGENNELPLTLEAVEPCSAHRRYFLFGSGVALFENTNFVKNWDAADIIQIATSTVEL 630
DB 623 EGQLEDNEIRLEONNLEPCTVGHKRYFVFGDGYVEFYEAYSHQVSRADVPVVFVDL 682
QY 631 NLTLLEDREILPLSVYTKELRDGVLDYAEVARNOLHEIK 672

DB 683 NLTMLEDHEFLPLEVYTRQEIKDGLDLYAEVORRNQMHAR 724
RESULT 9
VGLB_HSVE1 STANDARD; PRT; 980 AA.
AC P18050;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10327;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89279217; PubMed=2543744;
RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
RA Woodworth L.M.;
RA "Identification and nucleotide sequence of a gene in equine
RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
RT major envelope glycoprotein gB";
RL J. Gen. Virol. 70:383-394(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: D00401; BAA00304.1; ALT_SEQ.
DR PIR: A31241; VGBE2H.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86 POTENTIAL.
FT CHAIN 87 980 GLYCOPROTEIN B.
FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 853 870 POTENTIAL.
FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 980 AA; 109931 MW; 94B69AF882FB6AC2 CRC64;

Query Match 48.4%; Score 1697.5; DB 1; Length 980;
Best Local Similarity 47.3%; Pred. No. 1.9e-116;
Matches 338; Conservative 114; Mismatches 191; Indels 71; Gaps 13;

QY 19 TNSPSTQNTVS-----REVSSVOLSEESTFYLCPPPPGVST 56
DB 88 TTPSPPTSTSTHSGTVDTPLLPETDPLRLAVRESGILAEDEGDFVTCPPPTGST 147
QY 57 VIRLEPPKPCPEPRKATWEGGAILFKENISPKYKFKVLYKNIIOITTTGGTYTQIT 116
DB 148 WRLEPPTCPKFDGLRNFTEGIAVIFKENIAPYKFRANVYKRDIVVTVRWKGYSHTSLS 207
QY 117 NRYTDRPVSTEEITDIDGKRCSSKARYLRNNVYAEFDRDAGEKQV-LLKPSKENTP 175
DB 208 DRYNDRPVSVSEELFGLIDSGKCSKAEYLRDNIHMHAYHDDDEVELDLCRPS-LQLR 266

CC 176 ESRWHTTNET--YTWGSPWIRYRTGTSVNCIVEEMDARSVPYPSYFAMANGDIANISPF 233
 CC 267 GARAWOTINDTTSYVGM-NPWRHYTSTVNCIVEEARSVPYPSYFAMANGDIANISPF 325
 CC 234 YGLSPPEAAA--EPMGPODNFKQDLSYFMDLDRKRRKASLPVKRNFLLTSHFTVGDWA 291
 CC 326 YGL---RAAARTEHNSYAQERQVEGYRDLDSKLQAEPEVTKNFTHFTVTSWNT 382
 CC 292 PKTTRVCSMTKKEVTEMLRATVNGRYRFARELSATFISNTTEPDNRILGQCIRKA 351
 CC 383 EKKVEACTLTWKKEVDELVRDFRGSYRFTIRSISSFTISNTOPKLESAPLTCVSKEA 442
 CC 352 EAAIQIPRTKYNDSHVGVHGVQVFLALGGFTVAYOPVLSKSLAHMVLRLMRNRDTEM 411
 CC 443 KEADISYKKQYESTHVSFVGLYARGGFLAERPMLSNELARLYLNLVRSNRTYDL 502
 CC 412 LDVNNKHAIVKKNATSLRSLRDIRNAPNRKTIYD-----DTTA--- 451
 CC 503 KNLNPN-----ANNNTTTRRRSLLSVPEQPTQDGVHREQLHLRLKRAVEATACTDS 558
 CC 452 -----IKSTSSVOFAMQLFYDHIQTHINDMFRIATATWCELONRELVLWHEGK 501
 CC 559 SNVTAQOLEIKTTSSTEFAMQLFYDHIQSHVNEMLSRATATWCTONKERTLWNEVK 618
 CC 502 INPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTLQNSMRVITSTNYCSRPLVLF 561
 CC 619 INPSAIYSATLDERVAARVLDVIAITHCAKIEG-NVYLQNSMRSDS-NTCYSRPPVIF 676
 CC 562 SY---GENOGNIQGLGENNELPTLEAVEPCSANHRRYFLGSGYALPENYFVWMDA 618
 CC 677 TITKANNRGSIQGLGENNELPTLEAVEPCSANHRRYFLGSGYALPENYFVWMDA 618
 CC 619 ADIQIASTFVELNLTLEIREILPLSVYKTELDRGVLDYAEVARNOLHELK 672
 CC 737 TELEVISTVVELNLTLEIREILPLSVYKTELDRGVLDYAEVARNOLHELK 790

RESULT 10
 VGLB_HSV2H STANDARD; PRT; 904 AA.
 AC P08666; P89450;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN GB OR UL27.
 OS Herpes simplex virus (type 2 / strain HG52).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OX Alphaherpesvirinae; Simplexvirus.
 RN NCBI_TaxID=10315;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87071654; PubMed=3024391;
 RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
 RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
 RL comparison with the corresponding gene of HSV-1.";
 RN Virology 155:322-333(1986).
 RP SEQUENCE FROM N.A.
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: GH, GC, GD, GI, AND GE.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; M14923; AAA66440.1; -;
 CC EMBL; Z86099; CAB06752.1; -;
 CC InterPro; IPR000234; Glycoprot_B.
 CC Pfam; PF00606; Glycoprotein_B; 1.
 CC ProDom; PD006693; Glycoprot_B; 1.
 CC Glycoprotein; Transmembrane; Signal.
 CC SIGNAL 1 22
 CC CHAIN 23 904 GLYCOPROTEIN B.
 CC DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 728 743 POTENTIAL.
 CC TRANSMEM 749 768 POTENTIAL.
 CC TRANSMEM 772 792 POTENTIAL.
 CC DOMAIN 793 904 POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC N-LINKED (GLCNAC. .) (POTENTIAL).
 CC L -> V (IN REF. 1).
 CC T -> A (IN REF. 1).
 CC S -> T (IN REF. 1).
 CC L -> Q (IN REF. 1).
 CC S -> A (IN REF. 1).
 CC EL -> DV (IN REF. 1).
 CC R -> G (IN REF. 1).
 CC SEQUENCE 904 AA; 100217 MW; AB050A3APB4F1066 CRC64;
 Query Match 48.2%; Score 1693; DB 1; Length 904;
 Best Local Similarity 49.2%; Pred. No. 3.6e-116;
 Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;
 QY 20 NSSPSTQNTSREVSVSVQSEESTFYLCPPPVGVSTVIRLEPRKCEPEPKATWEGE 79
 DB 82 NATVAAGHATLRAHLREIKVENADAQAFVCPPTGATVQEPQPRCTRPEGQNYTEGI 141
 QY 80 AILFKENISPYKFKVLYKNIQITWTGTYTQITNRYDTRTPVSEIETDIDKGR 139
 DB 142 AVFENIAPYKFKATYIKDVTQSVQVFMGHRYSQFQIFEDRAPVPFEEVIDKNTKV 201
 QY 140 CSSKARYLRNNVYVEAFDRDAGEQVLLKPSKENTPESRAWHHTNEYTVWGSFVIYRTG 199
 DB 202 CRSTAKYVRNNMETTAFHRDDHETDMLKPAKATRTSRGWHTTDLKYNPSRVEAFHYG 261
 QY 200 TSVNCTIVEEMDARSVPYPSYFAMANGDIANISPFYGLSPPEAAEPQDQNFKOLDSY 259
 DB 262 TVNCTIVEEMDARSVPYPSYFAMANGDIANISPFYGLSPPEAAEPQDQNFKOLDSY 320
 QY 260 FMSDLKRRKASLPVKRNFLLTSHFTVGDWAPKTTRVCSMTKKEVTEMLRATVNGRYR 319
 DB 321 YARDLATTKARATSPTRNLLTTPKFTVANDWVPRPAVCTMTTKWQVDEMLRAEYGGSPR 380
 QY 320 EMARELSAIFISNTTEPDNRILGQCIRKREAAIEQIFRTKYNDSHVGVHGVQVFLAL 379
 DB 381 FSSDAISTTFTTNLTSEYLSRVDLDCIGRDAREADRMFAKYNATHKLVQGPQYILAT 440
 QY 380 GGFVAYQPVLSKSLAHMVLRLMRNRDTEMDLVNNKHAIVKKNATSLRDIRN 439
 DB 441 GGFVAYQPVLSKSLAHMVLRLMRNRDTEMDLVNNKHAIVKKNATSLRDIRN 482
 QY 440 PNKTILODTTAKSTSSVOFAMQLFYDHIQTHINDMFRIATATWCELONRELVLWHEG 499
 DB 483 PSANASVE---RIKTSSTIEFARLQFTYNNQRHVNDMLGRIVAWCMLQNLHETLWNEA 539
 QY 500 IKINPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTLQNSMRVITSTNYCSRPLV 559
 DB 540 RKLNPATASATVGRVRRSARMGLDMVAVSTCVVPADNVIVQNSRVSRRSGTCTSRPLV 599
 QY 560 LFSYGENOGN-IQGLGENNELPTLEAVEPCSANHRRYFLGSGYALPENYFVWMDA 618

```

Db 600 SFYR-EDQGLPGLGQNNELRLRDALPCTVGHRRYFIFGGGVYFEEYAYSHQLSR 658
QY 619 ADIQASTAVELNLTLEDEILPLSVYTKBELRDVGVDLYAEVARRNOLHELK 672
Db 659 ADVTIVSFIDLNITMLEDHEFVPLEYTRHEIKDSGLLDYTEVQRNQLDLR 712

RESULT 11
VGLB_HSV23
ID VGLB_HSV23 STANDARD; PRT; 904 AA.
AC P06763;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27 OR GB2.
OS Herpes simplex virus (type 2 / strain 333).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87112925; PubMed=3027364;
RA Stuve L.L., Brown-Shimer S., Pahl C., Najarian R., Dina D.,
RA Burke R.L.;
RT "Structure and expression of the herpes simplex virus type 2
RT glycoprotein gb gene.";
RL J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15118; AAA45837.1; -.
CC PIR; A26790; VGBEB2.
CC InterPro; IPR000234; Glycoprot_B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD000693; Glycoprot_B; 1.
CC Glycoprotein; Transmembrane; Signal.
CC SIGNAL 1 22 POTENTIAL.
CC CHAIN 23 904 GLYCOPROTEIN B.
CC DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 728 743 POTENTIAL.
CC TRANSMEM 749 768 POTENTIAL.
CC TRANSMEM 772 792 POTENTIAL.
CC DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 904 AA; 100186 MW; A8B36F74FDBC539 CRC64;

Query Match 48.2%; Score 1692; DB 1; Length 904;
Best Local Similarity 49.1%; Pred. No. 4.2e-116;
Matches 321; Conservative 115; Mismatches 194; Indels 24; Gaps 7;

QY 20 NSSPSTQNTVRSWSSVLSSEESTFYLCPPVGVSTVIRLEPPRCPCPRKATEWGEI 79
Db 82 NATVAAGHATLRAHLREIKVENADAQFYVCPPTGATVQVFQPRCPRPESQNTYEGI 141

```

```

QY 80 AILFENISPYKFKVLYYKNIQTTTGTGTYRQITNRYTDRTPVSIIEITDLIDGKR 139
Db 142 AVYFENIAPYFKATMYKYDVTQVWFGRHSYQFMGIFEDRAPVPFEEVIDKINAKV 201
QY 140 CSSKARYLANNVVEAFDRDAGEKQVLLKPSFENTPESRAWHTTNETYTVWGSPLRYNG 199
Db 202 CRSTAKYVRNNMETTAFHRDDHETDMELAPAKVATRTSGNWHHTDLUKYNPSVEAFHRYG 261
QY 200 TSVNCIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPEAAAAEPMGYPQDNFKOLDSY 259
Db 262 TTVNCIVEEDARSVPYDEFVLATGDFVYMSPFYGYR-EGSHTETSAAADRFKQVDGF 330
QY 260 FSWDLDKRRKASLPVKRNFLLTSHFTVGDWAPKTYRVCSTWKVETMLRATVNGRVR 319
Db 321 YARDLTTKARATSPTRNLLTTPKFTVADWVPKRPACTMTKWQEVDEMRAEYGGSR 380
QY 320 FMARELSATFISNTWFDPNRIILGOCIKREAAEQIFRTKYNDSHVKGHVQYFLAL 379
Db 381 FSSDALSTFTTNLTQYSLSRVDLGDICIGDAREADRMFAKYNATHIKVGQPYLAT 440
QY 380 GGFIVAYQVLSKSLAHMYLRELMRDNRDTEMLDLVNNKHAIYKKNATSLSLRDRINA 439
Db 441 GGLIAYQPLLNTLAEVYREVMEQ-----DRK-----PRNATP-----APREA 482
QY 440 PRKTYLDDTTAISTSSVQFAMQLYDHIQTHINDMFSTRATANCELORELVLWHEG 499
Db 483 PSANASVE---RIKTSSTIEFARLQFTYNNHQRHVNDMLGRIAVAVANCELOHETLWNEA 539
QY 500 IKINFSATASATLGRVAAKMGDVAASVSTDAIDAEVTLONSMEVITSTNCTYSRPLV 559
Db 540 RKLNPANIASATVGRVSRMGLDVMVAVSTCVPADNVIVQNSMRVSRPGTCYSRPLV 599
QY 560 LFSYENQGN-IGQGLGNNELPLTLEAVEPCSAHRRYFLFGSGYALFENFVKNVDA 618
Db 600 SFYR-EDQGLPGLGQNNELRLTRDALEPCTVGHRRYFIFGGGVYFEEYAYSHQLSR 658
QY 619 ADIQASTAVELNLTLEDEILPLSVYTKBELRDVGVDLYAEVARRNOLHELK 672
Db 659 ADVTIVSFIDLNITMLEDHEFVPLEYTRHEIKDSGLLDYTEVQRNQLDLR 712

RESULT 12
VGLB_HSVBC
ID VGLB_HSVBC STANDARD; PRT; 932 AA.
AC P12640;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein Gvp-6) (Glycoprotein I1A)
DE (Glycoprotein I6) (Glycoprotein G130) (Glycoprotein B).
GN GI OR UL27.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88300884; PubMed=2841484;
RA Whitbeck J.C., Bello L.J., Lawrence W.C.;
RT "Comparison of the bovine herpesvirus 1 gi gene and the herpes
RT simplex virus type 1 gb gene.";
RL J. Virol. 62:3319-3327(1988).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


EMBL; M21474; AAA46055.1; --
 DR EMBL; Z78205; CAB01598.1; --
 DR EMBL; AJ004801; CAA06106.1; --
 DR PIR; A28877; VGBERC.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 DR KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 67
 FT CHAIN 68 932 GLYCOPROTEIN I.
 FT DOMAIN 68 758 EXTRACELLULAR.
 FT TRANSXEM 759 827 POTENTIAL.
 FT DOMAIN 828 932 POTENTIAL.
 FT CARBOHYD 105 105 CYTOPLASMIC.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;
 Query Match 47.9%; Score 1681; DB 1; Length 932;
 Best Local Similarity 49.8%; Pred. No. 2.8e-115;
 Matches 330; Conservative 101; Mismatches 213; Indels 18; Gaps 6;
 QY 18 GTNSSPSTQNTVREVSVSSVQLSEESTFYLCPPVGVSTVIRLEPPKPCPEPRKATWGE 77
 DB 97 GDDAASPDNDTVRAALRLAQAAGENSREFFVCPGPGATVRLAPRCPPEYGLGRNYTE 156
 QY 78 GIALFKENISPKFKVLYKNIQTWTGTYTQINRYNDRTPVPSIEETDLIDGK 137
 DB 157 GIGVIKENIAPTFKAYI-YKNVIYTTWAGSTYAAITNQYTDVPGVNGEITDLVDRK 215
 QY 138 GRCSKARYLRNVVYAEADRDAGEKQVLLKPSKFWTPESRAWHHTNEYTYVWGSPIYR 197
 DB 216 WCLSKAEYLRGRKVAFDRDDDPWEAPLKLPAISGVRGWHHTDDVYALGSAGLYR 275
 QY 198 TGTSYNCIVEENDARSVPYSYFANAGDIANISPFYGLSPPEAAEPMPQDQFKOLD 257
 DB 276 TGTSYNCIVEEARSVPYDSEALSTGDIYMPFYGLR-EGAHREHTSPERFOQIE 334
 QY 258 SYFSDMLDKRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKWKVEVTEMLRATVNGR 317
 DB 335 GYIKRDMATGRRLKEPVSRLFRQHTVAVDWVPRKKNVCSLAKWREADEMLRDSRGN 394
 QY 318 YRFMARELSATFISNTTEFPDNRILLQCIKREAAEAEIQIFRTKYNDSHVKVGHVQYFL 377
 DB 395 FRTARSLATFYSDSHTFALQNVPLSDCVIEEAEAAVERVYRERNGTHVLSGSETYL 454
 QY 378 ALGGFIVAYQPVLSKSLAHMYRLRLMRDNTDEMOLDVNNKHAHYKKNATSLRLARDIR 437
 DB 455 ARGGFVAVFRPMLSNELAKIYLQELARSNGTLEGLFAA----AAPKGP---RRARAAP 507
 QY 438 NAPNRKITL-----DDTTAKTSSVQFAMQLQFLYDHIQTHINDMFRIATACELQN 490
 DB 508 SAPGGPGAANGPAGDGDAGGRVTVTSSAEAFALQTYDHIQTHVNTFMSRLATSWCLLN 567
 QY 491 RELVWHEGIIKNPSATASATILGRVAAKMLGDVAASVCSCTAIDAESVTLQNSMRVIST 550
 DB 568 KERALWAAKLNLPASAAALDRRAAAMLGDAVATYCHELGEGRVFIENSRR--APG 625
 QY 551 NTCYSRPLVLSYGENOCNIOGOLGNNELLPTLEAVEPCSANHRRYFLFGSGVALPENY 610
 DB 626 GVCYSRPLVSPAFNESEPEVGEQGEONELLPGRELVEPCPTANKHRYFRGADVYVYENY 685
 QY 611 NFVKWDAADIIQASTFVELNLTLEDREILPLSVYTKELRQVGLDYAEVARRNOLHE 670
 DB 686 AYRRVPLAEVISTVDNLAVLEDEFLPLEVTVRAELADTGLLDYSEIQRRNOLHE 745
 QY 671 LK 672
 DB 746 LR 747

RESULT 13
 VGLB_HSV2S
 ID VGLB_HSV2S STANDARD; PRT; 885 AA.
 AC P24994;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL27.
 OS Herpes simplex virus (type 2 / strain SA8) (Simian agent 8).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBTaxID=10316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B264;
 RX MEDLINE=91374035; PubMed=1895066;
 RA Borchers K., Weigelt W., Buhk H.-J., Ludwig H., Mankertz J.;
 RT "Conserved domains of glycoprotein B (gp) of the monkey virus, simian
 RL J. Gen. Virol. 72:2299-2304(1991).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: CH, GB, GC, GD, GI, AND GE.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 EMBL; X56935; CAA40256.1; --
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 DR KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 885 GLYCOPROTEIN B.
 FT DOMAIN 35 715 EXTRACELLULAR (POTENTIAL).
 FT TRANSXEM 716 731 POTENTIAL.
 FT TRANSXEM 737 756 POTENTIAL.
 FT TRANSXEM 760 780 POTENTIAL.
 FT DOMAIN 781 885 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 885 AA; 97811 MW; 39B495B329AB94E4 CRC64;
 Query Match 47.7%; Score 1674; DB 1; Length 885;
 Best Local Similarity 48.4%; Pred. No. 8.6e-115;
 Matches 319; Conservative 104; Mismatches 204; Indels 32; Gaps 4;
 QY 20 NSSPSTQNTVREVSVSSVQLSEESTFYLCPPVGVSTVIRLEPPKPCPEPRKATWGE 79
 DB 68 NASVEAGRATLREDLREIKARDGATFVCPVPGATVQFQPRCPADQGNVTEGI 127
 QY 80 AILFKENISPKFKVLYKNIQTWTGTYTQINRYNDRTPVPSIEETDLIDGK 139
 DB 128 AVVFKENIAPYKFAKATYIKDVTVSQVWFHGRYSQFMGIEDRAPVFEVMDKINAKV 187
 QY 140 CSSKARYLRNVVYAEADRDAGEKQVLLKPSKFWTPESRAWHHTNEYTYVWGSPIYR 199
 DB 188 CRSTAKYVRNNMESTAFHRDDHSDALAPAKAATRTSRGWHTTDLKYNPRAVEAPHRVG 247
 QY 200.TSVNCIVEEMDARSVPYSYFANAGDIANISPFYGLSPPEAAEPMPQDQFKOLD 259


```

Db 248 TTVNCIVVEEARSVYPIDEVLATGDFVYNSPPYGR-DGSHGEHAYAADRPQVDGY 306
Qy 260 FSDMLDKRRKASLPKRNFLITSHFTVGDWAPKTRVCSMTKWKVEYTEMLRAVNGRY 319
Db 307 YERDLSTGRRAAPVTRNLLTPKFTVGDWAPKRPSCVILTKWEVDEMLRAEYGPSFR 366
Qy 320 FMARELSATFISNTEFDPNRIILGQCICKREAAIEQIFRTKYNDSHVKVGHVQYFLAL 379
Db 367 FSSAALSTTEANRTEYALSVDLADCVGREAREAVDRIFLRRYNGTHVKVGQVQYLLAT 426
Qy 380 GGFIVAYOPVLKSLAHMYLRELRDNRDTEMLDVLNKNHAIYKKNATSLSLRDRINA 439
Db 427 GGFIIAQPLLSNADLYVRELVE-----QTRRPAAGD 461
Qy 440 PNRRIT----IDDTTA---IASTSVQFAMQLQFLYDHIQTHINDMFSRIATACELQNR 493
Db 462 PGEAATPGSPVDPVSVERIKTSSVEFARLQFTYDHIQRVNDMLGRIATACELQNR 521
Qy 494 VLWHEGKINPSATASATLGRRAAKMLGDVAVSSCTAIDAESVTQLQNSMRVITSTNWC 553
Db 522 TLMWNEARLNPGATASATVGRVRSARMLGDVMAVSTCVPAVDNVMQNSIGVAARPCTC 581
Qy 554 YSRPLVLFYSGENOGNIOGOLGENNELLPLEAVEPCSANHRRYFLGSGVALENYFV 613
Db 582 YSRPLVFRYBAGPLVEGQLEDNEIRLERDALEPCTVGHRRYFTFGAGVYFEFYAYS 641
Qy 614 KWDAAADIAQTASTVEINLTLEDREILPLSVYKKEELRDVGLDYAEVARRNOLHELK 672
Db 642 HOLGRADVTVSTINLTMLDEHFEVPLEYVTRQEKDGLDLYTEVQRNOLHALR 700

RESULT 14
VGLB_HSVB2
ID VGLB_HSVB2 STANDARD; PRT; 917 AA.
AC P12641;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B-1 precursor.
OS Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10296;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88306231; PubMed=2841793;
RA Hamerschmidt W., Contraths F., Mankertz J., Pauli G., Ludwig H.,
RA Buhk H.-J.;
RT "Conservation of a gene cluster including glycoprotein B in bovine
RL herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";
RN Virology 165:388-405(1988).
[2]
RP SEQUENCE OF 1-200 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hamerschmidt W., Contraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RL proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -1- FUNCTION: GBI IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE
CC PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A
CC SYNCTIAL PHENOTYPE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; M21628; AAA46052.1; -.
DR EMBL; M21632; AAA46052.1; -.
DR PIR; C29242; VGBEBH.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 917 GLYCOPROTEIN B-1.
FT TRANSMEM 578 594 POTENTIAL.
FT TRANSMEM 770 786 POTENTIAL.
FT TRANSMEM 795 811 POTENTIAL.
FT CARBOHYD 48 48
FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 917 AA; 101882 MW; 1B96CBF50DB4D3F3 CRC64;

Query Match 47.5%; Score 1666.5; DB 1; Length 917;
Best Local Similarity 48.1%; Pred. No. 3.2e-114;
Matches 314; Conservative 114; Mismatches 196; Indels 29; Gaps 3;

Qy 20 NSSPSTQNTSVREVSSVOLSEESTFVLCPPVGVSTVIRLEPRKCPKPKATEWGGI 79
Db 110 NASEPADPAELRADLRLKSSDDPNFVCPPTGATVVRLEPRCPKPELPGNFTGGI 169

Qy 80 AILFKENISPYKFKVLYKNIQTTWTGTYRQITNRYTDRTPVSEETITDLIDGKR 139
Db 170 AVTEKENLAPYKFKATMYKAVTVASVMSGYSYNQFNIPEDRAPIPEEIVDRHGRM 229

Qy 140 CSSKARLYRNYYEAFDRDAGERQVLLKPKFNPESRAHWTNETIYTVGSGWIYRTG 199
Db 230 CLSTAKYVRNNLETAFHNDADHEMKLVPAESAAPGLHRGWHHTRLKNNPTGSAWIHRHG 289

Qy 200 TSVNCIVVEEDARSVPVSYFAMANGDIANTSPFVGLSPPEAAAEPMGYPODNEKOLD 259
Db 290 TTVDCIVDEEAKSSYPNEFVATGDFVYASPPFGYR-DGSHSEHNAAYADRFQVDGF 348

Qy 260 FSDMLDKRRKASLPKRNFLITSHFTVGDWAPKTRVCSMTKWKVEYTEMLRAVNGRY 319
Db 349 FPRDFGTGRRHGSPTVYNLLTPMFTVGNWNAKRPSCVCTMTKWREVPMLRAEYGSFR 408

Qy 320 FMARELSATFISNTEFDPNRIILGQCICKREAAIEQIFRTKYNDSHVKVGHVQYFLAL 379
Db 409 FTSNALSATFTNLTQYSLSRVDLDCVGREAREADRIYLEKYNNHTHLRYGVSQYLLAT 468

Qy 380 GGFIVAYOPVLKSLAHMYLRELRDNRDTEMLDVLNKNHAIYKKNATSLSLRDRINA 439
Db 469 GGFIIAQPLLSNADLYVRELVE-----KPEERKLNAT----- 508

Qy 440 PNRRITLDDTTAISTSVQFAMQLQFLYDHIQTHINDMFSRIATACELQNRLEVLWHEG 499
Db 509 -----DGKVIITTSVEFARLQFTYVNHQKHNEMFGMAVSWCQLQNELTLWNEA 560

Qy 500 IKINPSATASATLGRRAAKMLGDVAVSSCTAIDAESVTQLQNSMRVITSTNCTYSLRV 559
Db 561 KKNIPASATVTLHRRYSACMLGDVLAISTCVAPPAENVIMQNSIRPKPTCYSRPLL 620

Qy 560 LFSYGENOGNIOGOLGENNELLPLEAVEPCSANHRRYFLGSGVALENYFVMDAA 619
Db 621 SFKHVDEELMEGOLGENNELRLDRDAVEPCSVGHKRYFLFGAGVYFEETYSHQLSRS 680

Qy 620 DIQIASTFVELNLTLEDREILPLSVYKKEELRDVGLDYAEVARRNOLHELK 672
Db 681 DITAVSTFIDLTMLDEHFEVPLEYVTRQEKDGLDLYTEVQRNOLHALR 733
```

```

RESULT 15
VGLB_HSVLF STANDARD; PRT; 903 AA.
ID P06436;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain F).
OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10304;
RN [1]
RN MEDLINE=85083254; PubMed=2981343;
RA Pellet P.E., Kousoulas K.G., Pereira L., Roizman B.;
RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
RT primary sequence and predicted protein structure of the wild type and
RT of monoclonal antibody-resistant mutants."
RL J. Virol. 53:243-253(1985).
RN [2]
RN SEQUENCE OF 1-176 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hamerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RL virology 185:406-418(1988).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14164; AAA45776.1;
DR EMBL; M21633; AAA45788.1;
DR PIR; A03750; VGBEB1.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 29
FT CHAIN 30 903 GLYCOPROTEIN B.
FT DOMAIN 31 729 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 730 745 POTENTIAL.
FT TRANSMEM 751 770 POTENTIAL.
FT TRANSMEM 774 794 POTENTIAL.
FT DOMAIN 795 903 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 903 AA; 100104 MW; 73BDCA7813DB35E8 CRC64;
Query Match 47.4%; Score 1663; DB 1; Length 903;
Best Local Similarity 47.6%; Pred No. 5.6e-114;
Matches 315; Conservative 110; Mismatches 199; Indels 38; Gaps 5;
QY 18 GTNSPQTQNTSREWSVQLSEESTFYLCPPVGVSTVIRLEPPKCPKPRKATEWGE 77
| :::: | | | :::: |::| |::| |::| |::| |::| |::| |::| |::|

```

```

Db 84 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVTVQFQEPQRCPTRPEGQNTYE 143
QY 78 GIALFKENISPYKFKYLYYKNLIQITTTTGTYYRQITNRYDRTPTVSIIEITDLIDGK 137
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 144 GIADVFKENTAPYKFKATMYKDVTSQVWFHGRYSQFMGIFEDRAPVPFEEVDKINAK 203
QY 138 GRCSKARYLRNNVYVEAFDRDAGEKOVLLKPKSFENTPESRAWHITTTTETVYVWGSPIYR 197
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 204 GVCSTAKYVYRNLETTFARHDDHETMELKPANAATRTSRGWHITTDLYKNPSRVEAFHR 263
QY 198 TGTSVNCIVEEMDARSVPFYSYFAMANGDIANTISPYGLSPPEAAAPMYPQDNFOLD 257
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 264 YGTVNCIVEEVDARSVPYDFVLAATGDFVYSPFFGYR-EGSHEHTSYAADREKQVD 322
QY 258 SYFSMDLDRKASLPVKNRFLTSHFTVGDWNAKPTTRVCSMTKKEVTEMLRATVNGR 317
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 323 GFVARDLTTKARATPTTRNLATPKFTVANDWVVKRPSVCTWTWKQEVDEMLRSEYGG 382
QY 318 YRFMARELSATFISNTTEFPNRIILGOCIKREAAIAIOIFRTKYNDSHVKVGHVQYFL 377
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 383 FRFSSDAISTTTTNTLTPSLSRVDLGDICGDARDMDRIFARRYNATHIKVGPQYYL 442
QY 378 ALGGFTVAYQPVLSKSLAHMYLRELMRDNTDEMLDLVNKKHAIYKKNATLSRLRDIR 437
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 443 ANGGLIAYQPLLSNTLAELYVRE-----HLREQSR 473
QY 438 NAPNRKITLDDTTA-----IKSTSVQFAMQLFLYDHIQTHINDMFSRIATAWCELQNR 491
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 474 KPPNTPPPPGASANASVERIKTTSIEFARLQFTYNHIQRHVNMDLGRVALWCELQNH 533
QY 492 ELVLNHEGKINPSATASATLGRVAAKMLGDAVAVSSCTADAESVTLONSMRVTSTN 551
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 534 ETLWNEARKLNPAIASATVGGRRVSARMLGDVMAVSTCVPAADNVIVQNSMRISRRPG 593
QY 552 TCYSRPLVLSYGENQGN-IOGQGENNELPTLEAVEPCSANHRRYFLFGSGVALFENY 610
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 594 ACYSRPLVSFRY-EDQGPLVEGOLGNNELRLTRDAIEPCTVGHRRYTFGGGYVYFEY 652
QY 611 NFVKMVAADIAQIASTFVELNLLEDEILPLSVYKKEELRDVGLDYAEVARNQLHE 670
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 653 AYSHQLSRADITVYSTFIDLNTMLEDBFVPLEVYTRHEIKDGLLDYTEVQRNQLHD 712
QY 671 LK 672
|:
Db 713 LR 714

```

Search completed: October 8, 2003, 17:03:42
Job time : 14.1001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:00 ; Search time 23.0198 Seconds
(without alignments)
2807.386 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score: 3510

Sequence: 1 MHYFRNCIFFLIVLYGTN.....DVGVLDYAEVARNOLHELK 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3510	100.0	865	1 VGBERB	glycoprotein B pre
2	1839	52.4	948	2 A56602	glycoprotein B hom
3	1735.5	49.4	980	1 VGBEC6	glycoprotein B pre
4	1726	49.2	868	1 VGBE31	glycoprotein B - h
5	1720.5	49.0	979	2 JH0109	glycoprotein 14 pr
6	1715.5	48.9	913	1 VGBEPS	glycoprotein gii p
7	1707.5	48.6	919	1 VGBEQH	glycoprotein B pre
8	1702.5	48.5	975	2 T42576	glycoprotein B pre
9	1697.5	48.4	980	1 VGBE2H	glycoprotein B pre
10	1696	48.3	920	1 D48349	glycoprotein B pre
11	1692	48.2	904	1 VGBEB2	glycoprotein B pre
12	1683.5	48.0	943	2 B48474	glycoprotein B - f
13	1681	47.9	932	1 VGBECB	glycoprotein gi pr
14	1674	47.7	885	1 VGBESA	glycoprotein B pre
15	1671	47.6	904	1 VGBEK2	glycoprotein B pre
16	1666.5	47.5	917	1 VGBEBH	glycoprotein B pre
17	1663	47.4	903	1 VGBEB1	glycoprotein B pre
18	1663	47.4	904	1 VGBEW7	glycoprotein B pre
19	1657	47.2	903	1 VGBEK1	glycoprotein B pre
20	1641	46.8	933	1 B48349	glycoprotein B pre
21	1549.5	44.1	928	1 VGBEBG	glycoprotein gi pr
22	1312	37.4	883	1 VGBEIS	glycoprotein B pre
23	1299.5	37.0	883	1 VGBEIL	glycoprotein B pre
24	1296.5	36.9	873	2 S26690	glycoprotein B - i
25	761	21.7	907	1 VGBETE	glycoprotein B pre
26	745	21.2	854	2 T03107	glycoprotein B - a
27	738.5	21.0	906	1 VGBEC1	glycoprotein B pre
28	737	21.0	928	1 VGBEMC	glycoprotein B pre
29	733	20.9	874	2 S25530	glycoprotein B - b

30	732	20.9	830	1 A44047	glycoprotein B pre
31	732	20.9	830	2 T43999	glycoprotein B [im
32	727	20.7	830	1 B44047	glycoprotein B pre
33	727	20.7	830	2 T44186	probable glycoprot
34	720.5	20.5	831	1 VGBE6S	glycoprotein B - h
35	716.5	20.4	822	2 T41941	glycoprotein B - h
36	694.5	19.8	857	1 QGBE1L	glycoprotein B - h
37	691	19.7	807	2 T42924	glycoprotein B - a
38	683.5	19.5	808	1 VGBESM	glycoprotein B pre
39	676	19.3	874	2 S55602	glycoprotein B - e
40	126.5	3.6	848	1 VGBED3	glycoprotein H pre
41	119	3.4	855	1 VGBE41	glycoprotein H pre
42	117.5	3.3	740	2 E81823	hypothetical prote
43	116	3.3	3415	2 A46105	hypothetical(NSI, N
44	115	3.3	1740	2 T43773	hypothetical prote
45	115	3.3	1824	2 T07589	disease resistance

ALIGNMENTS

RESULT 1

VGBERB

glycoprotein B precursor - Marek's disease virus (strain RB1B)

C:Species: Marek's disease virus

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C:Accession: A32402; B32402

R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

A>Title: Nucleotide sequence and characterization of the Marek's disease virus homolo

A:Reference number: A32402; M01D:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 <ROS>

A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271;304-330 <ROS2>

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-865/Product: glycoprotein B #status predicted <GPB>

F;709-728/Domain: transmembrane #status predicted <TM1>

F;732-752/Domain: transmembrane #status predicted <TM2>

F;27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 100.0%; Score 3510; DB 1; Length 865;

Best Local Similarity 100.0%; Pred. No. 1.4e-247;

Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSSPSTQNTSVTSREVSVSVQLSEESTFYLCPPPVGVSTVIRL 60

Db 1 MHYFRNCIFFLIVLYGTNSSPSTQNTSVTSREVSVSVQLSEESTFYLCPPPVGVSTVIRL 60

QY 61 EPPKCPKPKATKATWEGEGIAILFKENISPYKFKVLYYNNIIOTTWTGTYTQINRYT 120

Db 61 EPPKCPKPKATKATWEGEGIAILFKENISPYKFKVLYYNNIIOTTWTGTYTQINRYT 120

QY 121 DRTPVSEIETDLIDCKGRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPTESRAW 180

Db 121 DRTPVSEIETDLIDCKGRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPTESRAW 180

QY 181 HTTNETYTVMGSPWIYRTGTSVNCIVEEMDARSVPFYSYFAMANGDIANISPYGLSPPE 240

Db 181 HTTNETYTVMGSPWIYRTGTSVNCIVEEMDARSVPFYSYFAMANGDIANISPYGLSPPE 240

QY 241 AAAPNGYQDNFKQLDSDYFSMDLDRRRKASLPVKRNELITSHFTVGDWAPKTRVCMS 300

Db 241 AAAPNGYQDNFKQLDSDYFSMDLDRRRKASLPVKRNELITSHFTVGDWAPKTRVCMS 300

QY 301 TKWKEVTEMLRATVNGRYRFMARELSATFTSNTEFDPNRIILGQCICKREAEAAIQIFR 360

Db 301 TKWKEVTEMLRATVNGRYRFMARELSATFTSNTEFDPNRIILGQCICKREAEAAIQIFR 360

QY 361 TKYNDSHVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRDRTDEMLDLVNKNKHA 420
 |||||
 Db 361 TKYNDSHVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRDRTDEMLDLVNKNKHA 420
 |||||
 QY 421 IYKNATSLSLRLDIRNAPNRKITTLDTTAIAKSTSSVQFAMQLFDYDHIOTHINDMFSLR 480
 |||||
 Db 421 IYKNATSLSLRLDIRNAPNRKITTLDTTAIAKSTSSVQFAMQLFDYDHIOTHINDMFSLR 480
 |||||
 QY 481 IATAWCELQNLRELVLWHEGKINPISATASATLGRVAAKMLGDVAASVSSCTAIDAESVTL 540
 |||||
 Db 481 IATAWCELQNLRELVLWHEGKINPISATASATLGRVAAKMLGDVAASVSSCTAIDAESVTL 540
 |||||
 QY 541 QNSMRVITSTNCTSRPLVLSYGENOGNIQOGENNELPTLEAVEPCSANHRRYFLF 600
 |||||
 Db 541 QNSMRVITSTNCTSRPLVLSYGENOGNIQOGENNELPTLEAVEPCSANHRRYFLF 600
 |||||
 QY 601 GSGYALFENYFVRKMDAADIQIASTFVELNLTLLEDEILPLSYTKEELRDYGVLDYA 660
 |||||
 Db 601 GSGYALFENYFVRKMDAADIQIASTFVELNLTLLEDEILPLSYTKEELRDYGVLDYA 660
 |||||
 QY 661 EVARRNQLHELK 672
 |||||
 Db 661 EVARRNQLHELK 672
 |||||

RESULT 2

A56602
 glycoprotein B homolog precursor - feline herpesvirus 1
 C:Species: feline herpesvirus 1
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999
 C:Accession: A56602
 R:Maeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Niihara, M.; Kai,
 Arch. Virol. 127, 387-397, 1992
 A:Title: Identification and nucleotide sequence of a gene in feline herpesvirus type 1
 A:Reference number: A56602; MUID:93090104; PMID:1333759
 A:Accession: A56602
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-948 <MAE>
 A:Cross-references: GB:S49775; NID:g261094; PIDN:AB24381.1; PID:g261095
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:120002, NCBIP:120003)
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein

Query Match 52.4%; Score 1839; DB 2; Length 948;
 Best Local Similarity 51.7%; Pred. No. 1.1e-125;
 Matches 342; Conservative 126; Mismatches 186; Indels 8; Gaps 4;

QY 18 GTNSPSTQNTSREVYSSVQSLSEEE---STFYLCPPPGVSTVIRLEPPRKCPEPRKAT 73
 || |||
 Db 93 GTPPKPTDPTDMSDMREALRASQIEANGSPSTFYMCPPPGSGTVVRLEPPRACPDYKLGK 152
 |||||
 QY 74 EWGEGIALFKENTSPYKFKVLYKNIQTTTGTYYRQITNRYTDRTPVSTIEETDL 133
 |||||
 Db 153 NFTGIAVIPKENTIAPIKFRANIVYKNIIMTVWSSGYAVTNNRYTDRTPVVKVQETDL 212
 |||||
 QY 134 IDGRCSSKARYLRNYYVEAFDRDAGEQVLLKPKSFNTPESSRAWHNTTNETTVWGSP 193
 || |||
 Db 213 IDRGMLSKADYVYNNYQETAFDRDEDPRELPLKPKSFNTPPQSRGWHNTTNETYTKGAA 272
 |||||
 QY 194 WIYRTGTVNIVCEMDARSVPYYSYFAMANGDIANSIPYGLSPPEAAEPMGYPODNF 253
 |||||
 Db 273 GFHSGTGVNIVCEMDARSVPYYSYFAMANGDIANSIPYGLSPPEAAEPMGYPODNF 331
 |||||
 QY 254 QLDYSFMDLDRKASLPVKRNLFTISHTVGDWAPKTRVCSTKVKVEVEMLRAT 313
 |||||
 Db 332 QIBSGYPIDDLRLQLGAPVSRNFLETTPHYVAWNTTPKGRVCTLAKHREIDEMLRDE 391
 |||||
 QY 314 VNGRYRFARELSATFTSNTEPDNRIILGQCITKREAAIEQIFRTKYNDSHVKGHV 373
 |||||
 Db 392 YQGSYRTVTKTISATFISNTSQFEINRILGDCATKEAAEAIDRIYKSKYSTHIQTGTL 451
 |||||

QY 374 QYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRDRTDEMLDLVNKNKHAITYKKNATSL--SR 431
 :|||
 Db 452 ETYLARGGFLIAFRPMISNELAKLYINELARSNRTVDLSALLNPSETVQTRGSPVSNQ 511
 |||||
 QY 432 LRDIRNAPNRKI-TLDDTTAIAKSTSSVQFAMQLFDYDHIOTHINDMFSTRATACELQN 490
 |||||
 Db 512 HHRSSRTIEGGIETVNNASLLKTTSSVEFAMIQFAYDIQAHVNEMLSRITATACWCTLN 571
 |||||
 QY 491 RELVLWHEGKINPISATASATLGRVAAKMLGDVAASVSSCTAIDAESVTLQNSMRVITST 550
 |||||
 Db 572 REHLVWTETLKNLPGGVVSMALERRVSARLLGDVAVTQCVNISGHVYIIONSMRVTCSS 631
 |||||
 QY 551 NTCYSRPLVLSYGENOGNIQOGENNELPTLEAVEPCSANHRRYFLFGSGYALFENY 610
 |||||
 Db 632 TTCYSRPLVLSYGENOGNIQOGENNELPTLEAVEPCSANHRRYFLFGSGYALFENY 691
 |||||
 QY 611 NFVKWMDAADIQIASTFVELNLTLLEDEILPLSYTKEELRDYGVLDYAFAVARRNQLHE 670
 :|||
 Db 692 AYVRKVPVLSSELISAYVDNLTLLEDEILPLSYTKEELRDYGVLDYAFAVARRNQLHE 751
 |||||
 QY 671 LK 672
 ||
 Db 752 LK 753
 ||

RESULT 3

VGBC6
 glycoprotein B precursor - equine herpesvirus 1 (strain Ab4p)
 C:Species: equine herpesvirus 1
 A:Note: host Equus caballus (domestic horse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
 C:Accession: G36798
 R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
 submitted to GenBank, March 1992
 A:Description: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A36805
 A:Accession: G36798
 A:Molecule type: DNA
 A:Residues: 1-980 <TEL>
 A:Cross-references: GB:M86664; NID:g330791; PIDN:AAB02468.1; PID:g330825
 R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
 Virol. 189, 304-316, 1992
 A:Title: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A41831; MUID:92295566; PMID:1318606
 A:Contents: annotation; possible protein-coding frames
 A:Note: neither amino acid nor nucleotide sequence is given
 C:Genetics:
 A:Gene: 33
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-85/Domain: signal sequence #status predicted <SIG>
 F:67-78/Domain: transmembrane #status predicted <TM1>
 F:86-980/Product: glycoprotein B #status predicted <MAT>
 F:850-868/Domain: transmembrane #status predicted <TM2>
 F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (C

Query Match 49.4%; Score 1735.5; DB 1; Length 980;
 Best Local Similarity 48.0%; Pred. No. 3.9e-118;
 Matches 342; Conservative 113; Mismatches 189; Indels 69; Gaps 11;

QY 19 TNSPSTQNTS-----REVYSSVQSLSEESTFYLCPPPGVST 56
 |||||
 Db 88 TTPSPPTSTPTSMTHSGTVDPDLLPTDPLRLAVREGSILAEDGDFYTCPPPTGST 147
 |||||
 QY 57 VIRLEPRKCPEPRKATGEWEGIALFKENTSPYKFKVLYKNIQTTTGTYYRQIT 116
 |||||
 Db 148 VVREPTPCPKFDLGRNFTGIAVIFKENTAPKFRANIVYKNIQTTTGTYYRQIT 207
 |||||
 QY 117 NRYTDRTPVSTIEETDLIDGRCSSKARYLRNYYVEAFDRDAGEQVLLKPKSFNTPE 176
 |||||
 Db 208 DRYNDRVPSVEEIFGLIDSKGCSKAEYLRNIMHAYHDDDEVELDLVPSKFTATP 267
 |||||

Qy 177 SRAVHTNET--YTVWGSPWIRYRTGTSVNCIVEEMDARSVPYPYSFAMANGDIANISPPY 234
Db 268 ARAQITNDTTSYYGVW-MPWRYHTSTSYNCIVEEYVARSVPYDSFALLSTGDIYVASPPY 326
Qy 235 GLSPPEAAA--EPMGYPDQNFQSDSYFMDLDRKRSKASLPVKRRNFLTSTSHETVCGWDNP 292
Db 327 GL----RAAARTEHNSYAQERFQVEGYRPRLDQSKLQAEPEPTKFNITPHVTYSWNWTE 383
Qy 293 KTRVCSMTKWKEYTEMLRAVNGRYRFPARELSATFISNTTEFPDNRITILGQCICKRAE 352
Db 384 KKVEACTLTWKVEDELVRDEFGRSGYRPTIRSISSFTISNTQFKLESAPLTECVSKEAK 443
Qy 353 AARIEQIFPKYNDSHVKVGHVQYFLALGGFIVAQPVLSKLAHMYLRELMDRNTDEML 412
Db 444 EADISYKKQYESTHVFSGDVEYFLARGGFITAFRPLMSNELARLYLNLVRSNRYTDLK 503
Qy 413 DLVNKHAIYKKNATSISSLRRDIRNAPNRKITLD-----DTTA----- 451
Db 504 NLLNPN-----ANNNTTTRRRSSLLSVPEPQTQGVHREQLHLRHKRAVEATAGTSS 559
Qy 452 -----IKSTSVQFAMLOFYLDHIQTHINDMFSRIATACELQNRFLVWHEGKI 502
Db 560 NVTAKQLELIKTTSSIEFAMQFAYDHIOQSHVNEMLSRATACWCTLQNKERTLWNEMVKI 619
Qy 503 NPSATASATLGRYAAKMGDVAAVSSCTADAESVTLQNSMRVTSNTCYSRPLVLFS 562
Db 620 NPSAIVSNTLDERVAARVLGDVIAITHCAKTEG-NVYLQNSWRSDS-NTCYSRPPVTFT 677
Qy 563 Y---GENOGNTQGOLGENNELLPTLEAVEPCSAHHRYFLFGSGYALFENYFKWVDA 619
Db 678 ITKANNRGSTEGLQGENEFTETKRLKEPCALNQKRYFKGKVVYVYVETFRKVPPT 737
Qy 620 DIQIATFVELNLTLLDERTLPLSVYTKELRDVGVLDYAFVARNQLHELK 672
Db 738 EIEVISYVELNLTLLDEREFLPEVYTRAEELDGLDYSEIQRRNQHLAR 790

RESULT 4
VGBE31
glycoprotein B - human herpesvirus 3
N:Alternate names: glycoprotein II
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1998 #sequence_revision 30-Sep-1998 #text_change 16-Jul-1999
C:Accession: E27214
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: E27214
A:Molecule type: DNA
A:Residues: 1-868 <DAV>
A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27914.1; PID:G60020
C:Genetics:
A:Gene: 31
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:657-713/Domain: transmembrane #status predicted <TM1>
F:726-742/Domain: transmembrane #status predicted <TM2>
F:84,194,372,416,440,494,557,623,781,836,864/Binding site: carbohydrate (Asn)

	Query Match	49.2%	Score 1726;	DB 1;	Length 868;
	Best Local Similarity	50.5%	Pred. No. 1.6e-117;		
	Matches	325;	Conservative 116;	Mismatches 182;	Indels 20; Gaps 6
Qy	31	REVSSVQLSEESTFYLCPPVGSTVIRLPPKCPKPRKATWEGIAILFKENISPY	90		
		: : : : : : : :			
Db	41	REAIHKSQDAETKPTFYVCPPTGSTIVRLPFTTCDPYHLGKNFTGVIAVYKENIAAY	100		
Qy	91	KFKVLYAKNIQITTTGTVYRQITNRYTPTVSIIEIITDLIDGKRCSSKARYLRNN	150		
		: : : : : : : : : : :			
Db	101	KFKATVYIKDVIVSTAWAGSSITQITNRYARVPVPSIEITDIDFKGRCSSKATVYRNN	160		
Qy	151	VYVEAFDRDAGEKQVLKPKSPENTPESRAWHTTNTFTYVWGSPWLYRTGTGVNCEIVEEMD	210		

Db	161	HKVEAFNE	DKNQDMLPIASKYNSVSGKAWHTNDTYMVAGT	PGTGTGTSVNCIIIEE	220									
Qy	211	ARSVFYPSY	FAMANGDIANISPIYGLSP	PPAAAPGYP	PODNFKQJLDSFSDMLDRKKA	270								
Db	221	ARSIFPD	SFGLSTGDIYIYSPF	EGLR-DGAYREHSNI	AMDRPHQEGYRQRLDTRALL	279								
Qy	271	SLPKRNFLT	SFTVCGWDNAKPTTRVCS	MTWKKEVTEMLRATVNG	RYRFMARELSATFI	330								
Db	280	E-PAARNFL	YTPHLTVGVNWK	KPRTVC	SUVKRWREDDVVRDEYAHNFRFTKTLSTTFI	338								
Qy	331	SNTEFDP	NRILIGQCIKREAAE	AIQIFRTKYND	SHSVKVGHYQVYFALGSGFTVAYQPV	390								
Db	339	SETNEFN	LNQIHL	SQCVEEAR	LIINRIYTRYNS	SHVRTGDIQTYLARGGFFVVFQPLL	398							
Qy	391	SKSLAHMYL	RELMRDNETDEML	DLVNHKHA	IKKNATSLSR	LRDRINAPNRKITLDDTT	450							
Db	399	SNLSRLY	IQELVRE-----	NTNHS-PQKHPT	TRNTRSRSV-----	PVELRANR	441							
Qy	451	AIKSTSVQ	PAMQLFYLDHIQTH	INDMFSRIA	TAWCELO	NRELVLVHBEKINPSATASA	510							
Db	442	TIITTSV	FAMQLFYDHLQEH	VNEMLARIS	SSWCOLQNR	REALMSGLEFPINPSALAST	501							
Qy	511	TLGRRAAK	MLGDVA	VSSCTA	DAS-VTLQNS	MRVITSTNTCYSRPLV	SYGENQGN	569						
Db	502	ILDQVRK	ATILGDV	ISVSNCP	ELGSDTRI	LQNSMRVSGSTTRCYSRPLIS	VSLSNGSGT	561						
Qy	570	IQCOLGEN	NELLTLEAVE	PCSANHR	RYFLFG	SGYALFENYFN	VKAWDAADIOIASTFVE	629						
Db	562	VEGOLG	FDNLIMS	RLDLCE	PAHNRK	YFLFGHHY	YVYEDYRV	VREIAVDHVGMISTYVD	621					
Qy	630	LNITLLE	DRILPLSVY	TKELRD	GVLDV	AEAVARN	NOLHELK	672						
Db	622	LNITL	LKDRF	EPLOYITRE	LDRTGL	LDYSEI	QRNQMSLR	664						
RESULT	5													
Qy	JH0109	glycoprotein 14 precursor - equine herpesvirus 1												
C:	Species:	equine herpesvirus 1												
C:	Date:	31-Dec-1991	#sequence	revision	31-Dec-1991	#text	change	26-Aug-1999						
C:	Accession:	JH0109												
R:	Guo, P.													
Gene	87, 249-255, 1990													
A:	Title:	Characterization of the gene and an antigenic determinant of equine herpesv												
A:	Reference	number:	JH0109;	MUID:	90236317;	PMID:	192002							
A:	Accession:	JH0109												
A:	Molecule	type:	DNA											
A:	Residues:	1-979	<GEO>											
A:	Cross	references:	GB:M34861;	NID:	g330900;	PIDN:	AAA46086.1;	PID:	g330901					
A:	Note:	glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies,												
C:	Genetics:													
A:	Gene:	gpi4												
C:	Superfamily:	herpesvirus glycoprotein B												
C:	Keywords:	glycoprotein; transmembrane protein												
F:	1-86/Domain:	signal sequence	#status	predicted	<SIG>									
F:	87-979/Product:	glycoprotein 14	#status	predicted	<MAT>									
F:	831-871/Region:	hydrophobic												
F:	155,275,380,423,497,514,560,727,749,951,970/Binding	site: carbohydrate (Asn) (cova)												
Query	Match	49.0%;	Score	1720.5;	DB	2;	Length	979;						
Best	Local Similarity	47.7%;	Pred.	No.	4.9e-117;									
Matches	340;	Conservative	113;	Mismatches	191;	Indels	69;	Gaps	11;					
Qy	19	TNSPSTQNTS-----	-----	REV	VSVQLSEEST	FYLCPP	PVGST	56						
Db	88	TTPSPPTSTPTSM	STSHGTVDP	TLLPT	ETPDLRLA	VRESGILAE	DGFYTCP	PPGTGST	147					
Qy	57	VIRLEPRK	CPPEPRKAT	EWEGTAIL	FKENISY	KKVLY	YKNIQT	TWTGTT	116					
Db	148	VWRIEPT	CPKFDL	GRNFTE	GVIAVFK	ENIAPFR	ANYYKID	VYVYRW	MGYSTLS	207				
Qy	117	NRVTRDP	VPVSEIEET	DLIDG	KRCSSK	KARVLR	NNVYV	AFDRD	AGEKOVLLK	PKSNTPE	176			

Qy	177	SRAWHTTNET--YVWGSPNIYRTGTSVNCIVEEMDARSVPYSYFAMANGDITANISPFY	234
Db	268	ARAWQTNDATSYGV--MPWRHVTSTSVNCIVEEYARSVPYDSFALSTGDIYASPFY	326
Qy	235	GLSPPEAAA--EPMGYPQDNFKQLDYSFMDLDRKRRKASLPVARNFLITSHFTVGDWAP	292
Db	327	GL---RAAARIEHSYAQERFQVGYRPRDLDSKLQAEEPVTKNFITTPHVTVSNWTE	383
Qy	293	KTRVCSMTKWKVTEMLRATVNGRYRFMARELSATFISNTEFDPNRIILGOCIKREAE	352
Db	384	KKVEACTLTWKVEDELVRDFRGSYRFTTRSSSFISNTQFKLESAPLTECVSKEAK	443
Qy	353	AAIEQIFRTKYNDKSHVKGHVQVFLAAGGFIVAYQPVLSKLAHMYLRELMRDRNTDEML	412
Db	444	EADISYKKQYESTHVSFGDVEYTLARGGFIAFRPMLSNELARLYLNLVRSNRTYDLK	503
Qy	413	DLVNNKHAIYKKNATSLSLRRDIRNAPNKIILD-----DTA-----	451
Db	504	NLLNPN-----ANNNNNTTRRRSLLSVPEQPTQDGVHREQILHLRHKRAVEATAGDSS	559
Qy	452	-----IKSTSSVOFAMQLFDYHDIOHINDMFSAWCELNQRELVLWHEGIKI	502
Db	560	NVTAKOLELIKTTSSIEFAMQLQFAYDHIQSHVNEMLSRIATAWCTLQNKERTLNENVKI	619
Qy	503	NPSATASATLGRRYAAKMLGDVAASVCTAIDAESVTQLNSMRVITSTNTCYSRPLVFS	562
Db	620	NPSAIVSATLDERVAARVLGDVIAITHCAKIEG--NVYLQNSMRSDS--NTCYSRPPVFT	677
Qy	563	Y---GENOGNITOGLOGENNELLTLEAVEPCSAHHRFYFLFGSGYALFENYVFKWYDAA	619
Db	678	ITKNANNRGSGTEGOLGENSEIFTERKLIIEPCALNQKRYFKGEYVYVYENTFVRKVPT	737
Qy	620	DIQIASTFVELNLTLLDRILPLSVYTKELRDVGVLDYAEVARRNOLHELK	672
Db	738	EIEVISYVELNLTLLDREFLPLEVYTRAELEDTLGLDISEYQRRNLHALR	790

RESULT 4

VGBE31

glycoprotein B - human herpesvirus 3

N:Alternate names: glycoprotein II

C:Species: human herpesvirus 3, varicella-zoster virus

C>Date: 30-Sep-1998 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999

C:Accession: E27214

R:Davidson, A.J.; Scott, J.E.

J. Gen. Virol. 67, 1759-1816, 1986

A:Title: The complete DNA sequence of varicella-zoster virus.

A:Reference number: A27345; MUID:86306657; PMID:3018124

A:Accession: E27214

A:Molecule type: DNA

A:Residues: 1-868 <DAV>

A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27914.1; PID:g60020

C:Genetics:

A:Gene: 31

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:697-713/Domain: transmembrane #status predicted <TM1>

F:726-742/Domain: transmembrane #status predicted <TM2>

F:84,194,372,416,440,494,557,623,781,836,864/Binding site: carbohydrate (Asn) (covalent)

Qy	31	REVSVSVQLSEESTFYLCPPPVGSGTVIRLEPPKCPPEPRKATWEGGIIALPENISPY	90
Db	41	REAIHKSQDAETKPTFFVYCPPTGSGTIVRLEPTNCPDYLHKGKNTFEGIAVVKENIAAY	100
Qy	91	KFKVLYLYKNIQITWTGTYTQITNRYTDRTPVSTEEITDLIDGKRCSSKARYLRNN	150
Db	101	KFRATVYKDVYSTAWAGSSYQITNRYADRPVPIVSEITDIDKFGKCSKATYVRNN	160
Qy	151	VYVEAFORDAGEQVLLKPSKFNTPESRAWHHTTNETYTVWGSPWIYRTGTGTSVNCIVEEMD	210

Query Match 49.2%; Score 1726; DB 1; Length 868;

Best Local Similarity 50.5%; Pred. No. 1.6e-117;

Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;

```
Db 208 DRANDRPVSVVEELFLGLIDSGKCSKAEYLRDNIHMHATHDDEDEVELDLVPSKATPG 267
Qy 177 SRAWHHTNET--YTVGSPWIIYRTGTSVNCIVEEMDARSYPFYYSIFAMANGDIANTISPFY 234
Db 268 ARAWQTNDITTSYGV-MPWRHYTSTVNCIVEEVEARSVYDYSPALSTGDIVIYASPFY 326
Qy 235 GLSPPEAAA--EPMPGYPQDNFKQLDSYFMDLKRKASLPVKNFLITSHFTVGVNDWAP 292
Db 327 GL---RAAAIRIENSHYAQERFQVEGYRPRDLDSKLQAEPEPVTKNFITTHPVTVSNWTE 383
Qy 293 KTRVCSMTKWEKTEMLRATVNGRYRFMARLSATFISNTTFDPNRIILGOCIKREAE 352
Db 384 KKEVACTLTWKWEDELVRDFRGSYRFTTIRSISSYFISNTTFKLESAPLTCVSKK 443
Qy 353 AAIEQIPRTYRSHVKGHVQYFALGGFIVAYQPVLSLAHMYLRELNRDNRITDEML 412
Db 444 EADSIIYKKQYESTHVSFGVEYVYLAGGFLIAPRMLSNELARLYLNLVRSNRTYDLK 503
Qy 413 DLVNKKHAIYKKKATSLRRLRDRIRNAPNRKITLD-----DTTA----- 451
Db 504 NLLNPN----ANNNNNTTRRRSLLSVPEQPTQDGVRHQILHRLHKRAVEATAGTSS 559
Qy 452 -----IKSTSSVOFAMLOFLYDHIQTHINDMFSAIATAMCELQNRRELVLWHEGKI 502
Db 560 NVTAKQLELIKTTSSIEFAMLOFAYDHIQSHVNEMLSRITATACPLQNKERPLANEMVKI 619
Qy 503 NPSATASATLGRVAAKMLGDAVSSCTAIDAESVTLQNSMRVITSTNTCYRPLVLS 562
Db 620 TPSAIVSATLDERVAARVLGVDIAITHCAKIEG-NVYLQNSMRSDS-NTCYRPPVTF 677
Qy 563 Y---GENQNOGOLGNNELLPTLEAVEPCSANHRYFLFGSGYALFENYFNKVMYDAA 619
Db 678 ITKNANNRGSTEGOLGENEFTKRLIEPCALNQRKFKGYEYVYFNTFVRKVPPT 737
Qy 620 DIQASTFVELNLTLEDRELPLSVYTKBELRDVGVDYAEVARRNQLHELK 672
Db 738 EIEVISYVVELNLTLEDRELPLEVYTRAELEDTGLLDYSEIQRNQLHALR 790

RESULT 6
VGBEPS
glycoprotein gII precursor - suid herpesvirus 1
C:Species: suid herpesvirus 1
C:Date: 31-Mar-1998 #sequence_revision 31-Mar-1998 #text_change 16-Jul-1999
C:Accession: A29159
R:Robbins, A.K.; Dorney, D.J.; Wathen, M.W.; Whealy, M.E.; Gold, C.; Watson, R.J.; Holla
J. Virol. 61, 2691-2701, 1987
A:Title: The pseudorabies virus gII gene is closely related to the gB glycoprotein gene
A:Reference number: A29159; MUID:87284141; PMID:3039163
A:Accession: A29159
A:Molecule type: DNA
A:Residues: 1-913 <ROB>
A:Note: the authors translated the codon GAC for residue 860 as Asn
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-40/Domain: signal sequence #status predicted <SIG>
F:751-819/Domain: transmembrane #status predicted <TM>
F:820-913/Domain: intracellular #status predicted <INT>
F:151,261,441,516,633,697/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 48.9%; Score 1715.5; DB 1; Length 913;
Best Local Similarity 50.7%; Pred. No. 1e-116;
Matches 340; Conservative 105; Mismatches 184; Indels 41; Gaps 12;

Qy 18 GTNSSPSTQ--NVTSREVVSSVOLSEESTFLCPVPVGVSTVIRLEPPKPEPRKATEW 75
Db 95 GPSAPGEGYGLDARAVRAA--ATERDRYVCPPSGSTVIRLEPEQACPEYSQGRNF 152
Qy 76 GEGTALFKENISPYKVTLYYKNIITQTTWTGTTVQINRYTRDTPVSEIBTDLID 135
Db 153 TEGIAVLFKENIAPHKFAHYKNIIVTTWVSGSTVAAINRTDRVPVQEIITDID 212
```

```
Qy 136 GKRCSSKARYLNNVYVEAFDRDAGEKQVLLKPSKFNTPESPRAWHTTNETYTVWGSPI 195
Db 213 RRKCKYSKAEYVRNNHKVTAFRDENPVEYDLRPSRLNALGTRGWHNTTNDYTKGAAGF 272
Qy 196 YRTGTSVNCIVEEMDARSYPFYYSIFAMANGDIANTISPFYGLSPPEAAAEPMPGYPQDNFKQ 255
Db 273 YHTGTSVNCIVEEARSYPDYSPALSTGDIYVMSFFYGLR-EGAHGEHIGVAPGRFQ 331
Qy 256 LDSYFMDLKRKASLPVKNFLITSHFTVGVNDWAPKTRVCSMTKWKKEVTEMLR-ATV 314
Db 332 VEHYPIIDLDSRLRASESVTRNFLRTPHTVAMDWAPKTRRVCSLAKWREAEEMTRDETR 391
Qy 315 NGRYRFMARLSATFISNTTFDPNRIILGOCIKREAEAEIOIFRTKYNDSHVKVG-HV 373
Db 392 DGSFRFTSRALGASFVSDVTLQDLQRVHLGDCVLRSEADAIYRRRYNSTHVLGDRP 451
Qy 374 QYFALGGLFIVAYQPVLSLAHMYLRELNRDNRITDEMLDLVNKKHAIYKKKATSLR 433
Db 452 EYTLARGGFVAVRPLISNELAQIYAREL-----ERLGLAG---VGPAAAPAAARR 501
Qy 434 RD-----IRNAPNRKITLDITTAIKSTSSVOFAMLOFLYDHIQTHINDMFSAI 482
Db 502 RSPGAGTPEPPAVNGTGHRLIT-----TGSAEFARLQFTYDHIQAHVNDMLGRIA 552
Qy 483 TANCELQNRRELVLWHEGKIKNPSATASATLGRVAAKMLGDAVSSCTAIDAESVTLQ 542
Db 553 AAWCELQNKDRTLWSEMSRLNPSAVATAALGQVRSAARMGLGDMVAISRCVVRG-GVTVQN 611
Qy 543 SMRVITSTNTCYRPLVLSYGENQNOGOLGNNELLPTLEAVEPCSANHRYFLFGS 602
Db 612 SMRVPGERGTCYRPLVTFEH-NGTGVIEQLGDDNELLISROLIEPCTGNHRRYFLGS 670
Qy 603 GYALFENYFNKVMYDAAIDQIASTFVELNLTLEDRELPLSVYTKBELRDVGVDYAEV 662
Db 671 GYVYEDYNYVRMVEP--ETISTRTVLTNLTLEDRELPLEVYTRAELEDTGLLDYSEI 728
Qy 663 ARNQLHELK 672
Db 729 QRNQLHALK 738

RESULT 7
VGBEQH
glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C:Species: equine herpesvirus 4
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A31880
R:Riggio, M.P.; Cullinane, A.A.; Onions, D.E.
J. Virol. 63, 1123-1133, 1989
A:Title: Identification and nucleotide sequence of the glycoprotein gB gene of equine
A:Reference number: A31880; MUID:89125704; PMID:2915378
A:Accession: A31880
A:Molecule type: DNA
A:Residues: 1-919 <RIG>
A:Cross-references: GB:M26171; NID:g341446; PIDN:AAA46106.1; PID:g514920
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-919/Product: glycoprotein B #status predicted <GPB>
F:740-809/Domain: transmembrane #status predicted <TM>
F:106,216,321,364,438,456,493,499,666,688/Binding site: carbohydrate (Asn) (covalent)

Query Match 48.6%; Score 1707.5; DB 1; Length 919;
Best Local Similarity 47.9%; Pred. No. 3.9e-116;
Matches 336; Conservative 117; Mismatches 195; Indels 53; Gaps 11;

Qy 20 NSSPT-----QNTSREVVSSVOLSEES-----TFYLCPPPVGSGTVIRLE 61
Db 34 SSQSPSTASQSAKTVDQTLPTETPDPLRVAERSGILAEADGDFYCCPPPTGTVVRIE 93
Qy 62 PPRKCEPRKATEWEGEGIALFKENISPYKVTLYYKNIITQTTWTGTTVQINRYTRD 121
```

150	Db	: :: :: : :: : :	PPRSCPFDLGRNFTGEGIAVIFKENIAPYKFRANVYTKDVIIVTKVKGYSHTSLDRYND	209
122	QY	: : : :	RTVPSIEIEITDLIDGKRCSSKARYLRNNVVVEAFDRDAGSKOVLLKPSKFNTPSSRAWH	181
210	Db	: : : :	RVPVSVEIEITLIDSKGCSKAEYLRDNIHHAYHDEDEVELDVPKSKFATPGARAWQ	269
182	QY	: : : :	TTNET--YTVWGSPIVYRTGTSVNCVIEEMDARSVPFYSYFAMANGDIANISPFYGLSP	239
270	Db	: : : :	TTNDTTSYGVG-MEWRHYTSTSVNCVIEEVARSYVPDSEFALSTGDIIVTSPFYGLR-S	327
240	QY	: : : :	EAAPMGYPQDNFKQLDSDYSFMDLDRKRSKASLPVKRNFLLTSHFTVGDWDAKTRVCS	299
328	Db	: : : :	AAQLEHNSYAQERFQVEGYPRDLDSKLQAGEPYTKNFITPHVTVSWNTEKKIEACT	387
300	QY	: : : :	MTWKVEYTEMRLRVNGVGRPMARELSATFISNTTEFPDNPRIILGQCIKREAAEAIEQIF	359
388	Db	: : : :	LTKWKVEYDELVRDEFGRSYPRTIRSIISTFISNTQFKLEAPLTDCVSKAEKDAIDSII	447
360	QY	: : : :	RTKYNDSHVKGHVQYFLALGGFTVAYQVPLSKSLAHMYLRELMDRNTDEMILVNKKH	419
448	Db	: : : :	RQYESTHVSQGVFEFYLRGCFILAFRPMISNELARYLHELVRNSYNTIDUKNL-NPN	506
420	QY	: : : :	AIYKKNATS-----LSRLRRDIRNAPN----RKITLDDTTAKS	454
507	Db	: : : :	ANHNTNTRRSLLSIPEPTQESILHREQILHRLHKRAVEAANSTSSNVYAKOLELIKT	566
455	QY	: : : :	TSSVQFAMLOFLYDHIOTHINDMSRTATANCELRNRELVLWHEGIKNPSATASNTLGR	514
567	Db	: : : :	ASSTEFAMLOFAYDHIQSHVNMELSRITATACNTQNKERTLWNEMVKVNPSSAIVSATLDE	626
515	QY	: : : :	RVAAMGLDVAAVSCTAIDAESVTLQNSMRVITSTNCSRPLVLFYSYGEN---QGNIQ	571
627	Db	: : : :	RVAARVLGVDIAITHCVKIEG-NVYQLNSMR-SSDSNTCSRPPVFTFTTKNANSRGITIE	684
572	QY	: : : :	GOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKWYDAADIQTASFVELN	631
685	Db	: : : :	GOLGEENEVYTERKLIEPCALNQRIKFKGEIVYENITTVKRPVPTTEIEVISTVELN	744
632	QY	: : : :	LTLEDREILPLSVYTTKEELRDVGVDYAEVARNQLHELK	672
745	Db	: : : :	LTLEDREFPLEYVTRAELEDGTLGDYSEIORRNOLHALR	785

RESULT 9
VGB2EH
glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)
C/Species: equine herpesvirus 1
C/Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
C/Accession: A31241; JQ0008
R/Whalley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Ball, C.W.; Woodworth, L.
J. Gen. Virol. 70, 383-394, 1989
A/Title: Identification and nucleotide sequence of a gene in equine herpesvirus 1 ana
A/Reference number: A31241; MUID:89779217; PMID:2543744
A/Accession: A31241
A/Molecule type: DNA
A/Residues: 1-980 <WHA>
C/Superfamily: herpesvirus glycoprotein B
C/Keywords: glycoprotein; transmembrane protein
F/1-85/Domain: signal sequence #status predicted <SIG>
F/70-78/Domain: transmembrane #status predicted <TN>
F/86-980/Product: glycoprotein B #status predicted <MAT>
F/853-968/Domain: transmembrane #status predicted <FN2>
F/165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (C

```

Query Match      48.4%; Score 1697.5; DB 1; Length 980;
Best Local Similarity 47.3%; Pred.No.2.3e-115;
Matches 338; Conservative 114; Mismatches 191; Indels 71; Gaps 13;

F;165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (C

QY 19 TNSSPSTQWVTS-----REVWSVQLSEESTFVLCPPVGST 56
      ||| ||| |||
Db 88 TPSPPTSTPMSTHSGTVDPTLLPTETPDPLRLAVRESGILAEQDGYTCPPGTST 147

```



```
Qy 57 VIRLEPRKPEPRKATWEGEGTALFKENISPYKFKVLYYKNIIOITTTWTGTYRQIT 116
Db 148 WRIEPRTPCPLDGRNFTEGIAVIFKENIAPYKFRANVYKIDIVTVRWKGYSHTSLS 207
Qy 117 NRYTDRPVSEETDLDIGKRCSSKARLYRNRYVEAFDRDAGEKQV--LLRPSKENTP 175
Db 208 DRYNDRVPVSEEIFGLIDSGKSSKAEYLRDNIMHHAYHDDDEVELDLCRPS--LQLR 266
Qy 176 ESRAWHTTNET--YTVGSPWYRTGTVSNCIIVEMDARSVFPYSYFAMANGDIANTSPF 233
Db 267 GARAWQTNTDTISVGH--MWRHWTSTVNCIIVEARSVPYDSALSTGDIVIASPF 325
Qy 234 YGLSPPEAAA--EPMGYQDNFKOLDYSFMDLKRKASLPVKRNKLITSHFTVGGWMA 291
Db 326 YGL---BAAARIEHNSYAQERFQVGYRPRDLQKQAEPEVTKNFITTPHVTWSNWT 382
Qy 292 PKTRVCSMTKWKKEVTEMLRATVNGRYREMARELSATFISNTTEFDNRIILGOCIKREA 351
Db 383 EKKVEACTLTAKKEVDELVRDEFSGSYRFTIRSISSFTSINTQFKLESAPLTCVSKEA 442
Qy 352 EAAIEQIFRTKYNDSHVKGHVQVFLALGGFIVAYQVPLSKSLAHMYLRLMRDRNRTDEM 411
Db 443 KEAIDSIVKQYESTHVSFGDVEYLLARGGFIAFRPMLSNEARLYNLNVLNRSNRTYDL 502
Qy 412 LDLYNNKHAIYKKNATSLSLRLRRDIRNAPNKKITLD-----DYTA----- 451
Db 503 KNLLNPN-----ANNNTTTRRRSLLSVPPEQPTQDGVHREQLHRLHKKRAVEATAGTDS 558
Qy 452 -----IKTSSTVQFAMQLFYDHIQTHINDMFSRIATACELQNLRELVLWHEGIIK 501
Db 559 SNVTAKOLELKTTSSTIEFAMQLQFAYDHIQSHVNMELSRATANTCTQNKERTLWNEVK 618
Qy 502 INPSATASATLGRVAAKMGDVAASCTAIDAESVTLQNSMRVITSTNTCYSRPLVLF 561
Db 619 INPSAIVSATLDERVAARVLGDVIAITHCAKIEG--NVYTLQNSMRSDS--NTCYSRPPVTF 676
Qy 562 SY---GENOGNIQOLGNNELLTLEAVEPCSANHRRYFLFGSGYALFENYVKNVDA 618
Db 677 TITNANNRSGIEGOLGNEEFTERKLIEPCALNQKRYFKFYVYENYITFVRKVP 736
Qy 619 ADIQTASTFVLENTLLEDRILPLSVYTKBELRDVGVLDAEYARRNQLHELK 672
Db 737 TEIEVISTVYELNLTLEDRFLPLEVYTRAELEDTLGLDYSEIQRRNQLHALR 790

RESULT 10
D48349
glycoprotein B precursor - saimiriine herpesvirus 1 (strain MV-5-4-PSL)
C:Species: saimiriine herpesvirus 1
C:Date: 17-Feb-1994 #sequence_revision 1
C:Accession: D48349
R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A:Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine m
A:Reference number: A48349; MUID:93228440; PMID:8385913
A:Accession: D48349
A:Molecule type: DNA
A:Residues: 1-920 <EB>
A:Note: sequence extracted from NCBI backbone (NCBIN:129066, NCBIPI:129068)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-920/Product: glycoprotein B #status predicted <MAT>
F:576-592/Region: hydrophobic
F:761-780/Domain: transmembrane #status predicted <TM1>
F:787-804/Domain: transmembrane #status predicted <TM2>
F:98,119,152,409,441,683,733/Binding site: carbohydrate (Asn) (covalent) #status predict
Query Match 48.3%; Score 1696; DB 1; Length 920;
Best Local Similarity 49.3%; Pred. No. 2.7e-115;
Matches 322; Conservative 95; Mismatches 210; Indels 26; Gaps 3;
Qy 20 NSSPSTQNTVSREVSSVQLSEESTFYLCPPVPGVSTVIRLEPPKPEPRKATWEGSGI 79
```

```
Db 98 NKTAEDARAQLRSVRQIRAEANATSMFYVCPPTGATVQVFEPRPCPDVAAGKNFTEGI 157
Qy 80 AILFKENISPYKFKVLYYKNIIOITTTWTGTYRQITNRYTTPVSEIEETDLDIGKGR 139
Db 158 AVIFKENIAPYKFKVLYYKNIIOITTTWTGTYRQITNRYTTPVSEIEETDLDIGKGR 217
Qy 140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSKENTPESRAWHTTNETYTVWGSPIWYRTG 199
Db 218 CRSDVTYVRSORRYTAYDRDGEVGLVPSKTSPTNSRGWYTTDRMYPANAHAGFYKAG 277
Qy 200 TSVNCIIVEMDARSVFPYSYFAMANGDIANTSPYGLSPPEAAAEPMGYPQDNFKOLDYSY 259
Db 278 TTVNCIIVEARSAYPYSNFVLTATGDFVYVSPYGLG--EDAHEYNAYSADREFKQVDGF 336
Qy 260 FSDMLDRRRKASLPVKRNKLITSHFTVGDWKPATKTRVCSMTKWKKEVTEMLRATVNGRYR 319
Db 337 FPRDLDSGETAPEVPVRNLLTTPKFTIGDWKPKDPSVCSVTKWEVEEMRAEYGSTFR 396
Qy 320 FMARELSATFISNTTEFDNRIILGQCIIKRAEAAIQIFRTKYNDSHVKGHVQYFLAL 379
Db 397 FTSSLSATFTTNTVTPPORIELSDCVAREQAADVAIYARRYNASHVKVGLQYTLAQ 456
Qy 380 GGFTVAYQVPLSKSLAHMYLRLMRDRNRTDEMLDLVNKHAIYKKNATSLSLRLRRDIRNA 439
Db 457 GGLVTVYQPLISLAEMYLRE-----AERALEPAPLPTTPA 494
Qy 440 PNRRITLDDTTAIRSTSVQFAMQLFYDHIQTHINDMFSRIATACELQNLRELVLWHEG 499
Db 495 PE---AAGSRGTLSTTQSVFARLQFTYDHIQKHVNMELGRIAAAWCOLQNOELVLWNEA 551
Qy 500 IKNPSTASATLGRVAAKMGDVAASCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 552 RKLNPATASATLGRVAAKMGDVAASCTAIDAESVTLQNSMRVITSTNTCYSRPLV 611
Qy 560 LFSYGENOGNIQOLGNNELLTLEAVEPCSANHRRYFLFGSGYALFENYVKNVDA 619
Db 612 SFRYADGELIEGOLGNEEFTERKLIEPCALNQKRYFKFYVYENYITFVRKVP 671
Qy 620 DIQTASTFVLENTLLEDRILPLSVYTKBELRDVGVLDAEYARRNQLHELK 672
Db 672 DVPVSTFVLDNLMTLEDRFLPLEVYTRAELEDTLGLDYSEIQRRNQLHALR 724

RESULT 11
VGEBE2
glycoprotein B precursor - human herpesvirus 2 (strain 333)
C:Species: human herpesvirus 2
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A26790; A45928
R:Stuve, L.L.; Brown-Shiner, S.; Pachl, C.; Najarian, R.; Dina, D.; Burke, R.L.
J. Virol. 61, 326-335, 1987
A:Title: Structure and expression of the herpes simplex virus type 2 glycoprotein gB
A:Reference number: A26790; MUID:87112925; PMID:3027364
A:Accession: A26790
A:Molecule type: DNA
A:Residues: 1-904 <STU>
A:Cross-references: GB:M15118; NID:g330256; PIDN:AAA45837.1; PID:g330257
R:Zwaagstra, J.C.; Leung, W.C.
Can. J. Microbiol. 33, 879-887, 1987
A:Title: The nucleotide sequence of herpes simplex virus type 2 (333) glycoprotein gB
A:Reference number: A45928; MUID:88079667; PMID:2446730
A:Accession: A45928
A:Molecule type: DNA
A:Residues: 1-34, 'AWPTV', 42-307, 'T', 309-481, 'R', 483-609, 'M', 611-664, 'R', 666-904 <ZWA>
A:Cross-references: GB:M24771; NID:g341245; PIDN:AAA60540.1; PID:g623400
A:Note: the authors translated the codon ATG for residue 610 as Ile
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAT>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
```


C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

;Molecule type: DNA
 ;Residues: 1-932 <WHI>
 ;Cross-references: GB:M21474; NID:g330756; PIDN:AAA46055.1; PID:g330757

R:van Drunen Littel-van den Hurk, S.; Parker, M.D.; Fitzpatrick, D.R.; van den Hurk, J.V.
 Virology 190, 378-392, 1992
 A:Title: Structural, functional, and immunological characterization of bovine herpesvirus
 A:Reference number: A43380; MUID:92410615; PMID:1326809
 A:Accession: A43380
 A:Molecule type: protein
 A:Residues: 68-87;505-516 <VAN>
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-67/Domain: signal sequence #status predicted <SIG>
 F:68-932/Product: glycoprotein gi #status predicted <GIB>
 F:68-504/Product: glycoprotein gib #status predicted <GIC>
 F:505-932/Product: glycoprotein gic #status predicted <GIC>
 F:759-827/Domain: transmembrane #status predicted <TM>
 F:105,153,441,483,640,706/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 47.94; Score 1681; DB 1; Length 932;
 Best Local Similarity 49.8; Pred. No. 3.4e-114;
 Matches 330; Conservative 101; Mismatches 213; Indels 18; Gaps 6;

QY 18 GTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPKCPPEPRKATEWGE 77
 DB 97 GDDAASPNSVDVRAALQLAQAQENSRFFVCPPPGATVVR LAPARCPCEYGLGRNYTE 156

QY 78 GTALFKENISPKFKVLYKNIQTWTGTYTQITNRYTDRTPVSIIEITDLIDGK 137
 DB 157 GIGVIYKENIAPYTFKAVI-YKNVLTWTWAGSYAAITNQYTDVVPVGMGEITDLVDKK 215

QY 138 GRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPEAWHTTNEYTVWGSPIYR 197
 DB 216 WRCLSKAEYLRSGRWAFDRDDQFWEAPLAPRLASAPGVGRWHTTDDVYVYALSGAGLR 275

QY 198 TGTSTVNCIVEEMDARSVPYFYFAMANGDIANSIPFYGLSPPEAAAEPMGYPDQNFOLD 257
 DB 276 TGTSTVNCIVEEARSVPYFSDALSTGDIYMSPFYGLR-EGAHREHTSYSPERFOQIE 334

QY 258 SYFSMDLDRKASLPVKNFLITSHFTVGDWAPKTTVCSMTKWKVEYTEMRLATVNGR 317
 DB 335 GYIKRDMATGRKKEPVSRNFLTQHTVAVDWVDPKRNVCSLAKWREADEMLRDESGN 394

QY 318 YRFMARELSATFISNTTEFPDNPRIILGQCIKRAEAAIEQIFRTKYNDSHVKVGHVQYFL 377
 DB 395 FRETARLSATFVSDSHTFALQNVPLSDCVIEEAAEAVRYRYNGTHVLSGSLTYL 454

QY 378 ALGGFTVAYQPVLSKSLAHMYLRELMNRDTEMDLNVNKHAIYKKNAATSLSLRRDIR 437
 DB 455 ARGGFVAPRPMLSNELAKLYLQELARSNGTLEGLFAA----AAPKPGP---RRARRAAP 507

QY 438 NAPNRKITL-----DOTTAIKSTSSVQFAMQLFYDHIOTHINDMFSTRATACELQ 490
 DB 508 SAPGGPGAANGPAGDAGGRVTTVSSAEFAALQFYDHIQHVTFMFSRLATSWCLLQ 567

QY 491 RELVWHEGIKINPSATATLGRVAAKMLGDVAASCTAIDAESVTQLQNSMRVITST 550
 DB 568 KERALWAEAKLNPSAASALDRRAARMGLDAMAVTYCHELGEGRVFIENSMR--APG 625

QY 551 NTCYSRPLVFSYGENQNIQOLGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENY 610
 DB 626 GVCYSRPPVSAFAGNESEVEQGLGDNELLCPRELVEPCTANHRRYFRFGADYVYENY 685

QY 611 NFVKWDAADIOASTFVFNLTLEDEILPLSVYTKELROVGLDYAEVARNQLHEK 670
 DB 686 AYVRRVPLAELEIVSTFVDNLNLTLEDEFLPFLVYTRAEALDTGLLDYSEIQRRNLHE 745

QY 671 LK 672
 DB 746 LR 747

RESULT 14
 VGBESA
 glycoprotein B precursor - simian herpesvirus SA8 (strain B264)
 C:Species: simian herpesvirus SA8

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
 C:Accession: JQ1332
 R:Borchers, K.; Weigelt, W.; Buhk, H.J.; Ludwig, H.; Mankertz, J.
 J. Gen. Virol. 72, 2295-2304, 1991
 A:Title: Conserved domains of glycoprotein B (gB) of the monkey virus, simian agent 8
 A:Reference number: JQ1332; MUID:91374035; PMID:1895066
 A:Accession: JQ1332
 A:Molecule type: DNA
 A:Residues: 1-885 <BOR>
 A:Cross-references: EMBL:X56935; NID:g60438; PIDN:CAA40256.1; PID:g60439
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-885/Product: glycoprotein B #status predicted <MAT>
 F:712-732/Domain: transmembrane #status predicted <TM1>
 F:737-752/Domain: transmembrane #status predicted <TM2>
 F:760-780/Domain: transmembrane #status predicted <TM3>
 F:68,122,379,411,659/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:97-538,114-514,188-252,345-393,581-618/Disulfide bonds: #status predicted

Query Match 47.74; Score 1674; DB 1; Length 885;
 Best Local Similarity 48.4; Pred. No. 1e-113;
 Matches 319; Conservative 104; Mismatches 204; Indels 32; Gaps 4;

QY 20 NSSPSTQNTVSREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPKCPPEPRKATEWGE 79
 DB 68 NASVEAGRATLREDLREIKARDGATFYCPTTGATVQVQEPQPCRPAPDQONTTEGI 127

QY 80 AILFKENISPKFKVLYKNIQTWTGTYTQITNRYTDRTPVSIIEITDLIDGKR 139
 DB 128 AVFKENIAPYTFKAVI-YKNVLTWTWAGSYAAITNQYTDVVPVGMGEITDLVDKK 187

QY 140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPEAWHTTNEYTVWGSPIYRTG 199
 DB 188 CRSTAKYVRNNMESTAFRRDDHSDMALPKAKAATRTSRGWHTTDLKYNARVEAPRYG 247

QY 200 TSVNCIVEEMDARSVPYFYFAMANGDIANSIPFYGLSPPEAAAEPMGYPDQNFOLD 259
 DB 248 TTVNCIVEEARSVPYFDEFVLATGDFVYMSPFYGR-DGSHGEHTAYAAADRFRQVDGY 306

QY 260 FMSDLDRKASLPVKNFLITSHFTVGDWAPKTTVCSMTKWKVEYTEMRLATVNGR 319
 DB 307 YERDLSTGRRAAPVTRNLTTPKFTVGDWAPKRPVSCITLTKWREDEMLRAEYGFSPR 366

QY 320 FMARELSATFISNTTEFPDNPRIILGQCIKRAEAAIEQIFRTKYNDSHVKVGHVQYFL 379
 DB 367 FSSAALSTFTFANTEVALSRVLDLADCVGRAREAVDRIFLRRYNGTHVKVQVQYVLT 426

QY 380 GGFTVAYQPVLSKSLAHMYLRELMNRDNTDEMLDVLNKHAIYKKNAATSLSLRRDIRNA 439
 DB 427 GGFLIAYQPLLSNALVELYVRELVR-----QTRRPAGGD 461

QY 440 PNKRKITL-----LDDTTA--IKSTSSVQFAMQLFYDHIOTHINDMFSTRATACELQ 493
 DB 462 FGEATPGSDPPSVVERIKTTSVSEFARLQFYDHIQHVNDMLGRATATACELQREL 521

QY 494 VLWHEGIKINPSATATLGRVAAKMLGDVAASCTAIDAESVTQLQNSMRVITSTNTC 553
 DB 522 TLWNEARLNPGATASAVGRVSRMGLDMNAVSTCVVPADPNVIMQNSIGVAARPGTC 581

QY 554 YSRPLVFSYGENQNIQOLGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENYFV 613
 DB 582 YSRPLVFSYFADGDLVEGQLGEDNEIRLEDALEPCTVGHRRYFTFGAGYVFEYAYS 641

QY 614 KMWDAADIOASTFVFNLTLEDEILPLSVYTKELROVGLDYAEVARNQLHEK 672
 DB 642 HOLGRADVTVTSTFINLNLTLDEHFLPFLVYTRAEALDTGLLDYSEIQRRNLHAR 700

RESULT 15
 VGBEK2
 glycoprotein B precursor - human herpesvirus 2 (strain HG52)
 C:Species: human herpesvirus 2

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A25611
R;Bzik, D.J.; Debroy, C.; Fox, B.A.; Pederson, N.E.; Person, S.
Virology 155, 322-333, 1986
A;Title: The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with
A;Reference number: A25611; MUID:87071654; PMID:3024391
A;Accession: A25611
A;Molecule type: DNA
A;Residues: 1-904 <BZI>
A;Cross-references: GB:M14923; NID:g330254; PIDN:AAA66440.1; PID:g330255
R;Norais, N.; Tang, D.; Kaur, S.; Chamberlain, S.H.; Maslars, F.R.; Burke, R.L.; Marcus,
J. Virol. 70, 7379-7387, 1996
A;Title: Disulfide bonds of herpes simplex virus type 2 glycoprotein gB.
A;Reference number: A58366; MUID:97048015; PMID:8892856
A;Contents: annotation; tryptic peptide disulfide bond assignments
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-904/Product: glycoprotein B #status predicted <WAT>
F;724-744/Domain: transmembrane #status predicted <TM1>
F;749-769/Domain: transmembrane #status predicted <TM2>
F;772-792/Domain: transmembrane #status predicted <TM3>
F;82,136,250,393,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status pre
F;111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status experimental

Query Match 47.5%; Score 1671; DB 1; Length 904;
Best Local Similarity 48.5%; Pred. No. 1.7e-113;
Matches 317; Conservative 117; Mismatches 196; Indels 24; Gaps 7;

QY 20 NSSPSTQNTVSRWSSVOLSEESTFYLCPPPGVGTIVLEPPKCPKPKATWEGEGL 79
DB 82 NATVAAGATVRAHLREIKVENADAFYVCPPTGATVQVQEPQRCPTPEQNTTEGI 141
QY 80 AILFKNISPKFKVLYKNIITQTWTGTYTQITNRYTDRTPVSIEITLIDKGR 139
DB 142 AVVEKENIAPYKFKATMYKDVTVSQWFGHYRYSQFMGIFEDRAPVPFEVIDKINAKGV 201
QY 140 CSSKARYLRNNVVEAFDRDAGEKQVLLKPKSKNTPESSRAWHHTNETHYVWGSFWIYRTG 199
DB 202 CRSTAKYVRNMETTAHEDDHDHETMELPAKVATRTSGWHTTDLKYNPSRYEAFHRYG 261
QY 200 TSYNCIVEENDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMPQDNFKOLDY 259
DB 262 TTVNCIVEENDARSVPYDEFLVATGDFVYMSPFYGR-EGSHTHTTAADRPFKQVDGF 320
QY 260 FMDLDRKRKASLPVKRNLITSHFTVGMWAPKTRVCSMTKWEVTEMLRATVNGRYR 319
DB 321 YARDLTTKARATSPITRNLLTTPKFTVANDVPKRPVACTMTKQVDEMLRAEYGSFR 380
QY 320 FMARELSATFISNTTFDPNRIILGOCIKREAAAEQIFRTKYNSHVKGHVQYFLAL 379
DB 381 FSSDAISTFTTTLNTEYSLSRVDLGDICGRDAREADRMFARKYNATHIKVGPQYQAT 440
QY 380 GGFIVAYQVLSKSLAHMYLRELDRNRDTEMLDVLNNKHAIYKKNATSLSLRDRINA 439
DB 441 GGFIIAYQVLSWFLAEIYREYMQE-----DRK-----PRNATP-----APLEA 482
QY 440 PNKRITLDDTTAKTSSVQFAMQLQFLYDHIQTHINDMFESRIATAMCELQNLRLVWHEG 499
DB 483 PSANASVE---RIKTTSSIEFARLQFTYNHQIRHVNDMLGRIAVANCELQNHETLWNEA 539
QY 500 IKINPSATASATLGRVAAKMGDAVAVSCTAIDAESVTQLNSMRVITSTNTCYSRPLV 559
DB 540 RKLNPNAIASATVGRVRSARMLGDVMAVATCPVAPDNVIVQNSMRVSSRPGTCYSRPLV 599
QY 560 LFSYGENOGR-IOGOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVA 618
DB 600 SFYR-EDQGPLEIEQUGENNDVRLTRDALEPCTVGHNGYFIFGGGYVYFEEYASHQLSR 658
QY 619 ADIQIASTFVELNLTLEDREILPLSVYTKELRDVGLDYAEVARRNQLHELK 672
DB 659 ADVTTVSTFDLNTMLEDHEFVPLEVYTHREIKDSGLLDYTEVQRNQLHDLR 712

Search completed: October 8, 2003, 17:08:31
Job time : 26.0198 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:50 ; Search time 12.6904 Seconds
(without alignments)
1552.587 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score: 3510

Sequence: 1 MHYFRNCIFPLVILYGTN.....DVGVLDAEVAARNQLHELK 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 137520 seqs, 29319821 residues

Total number of hits satisfying chosen parameters: 137520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA.New.*
1: /cgn2.6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
2: /cgn2.6/ptodata/2/paa/US06_NEW_COMB.pcp.*
3: /cgn2.6/ptodata/2/paa/US07_NEW_COMB.pcp.*
4: /cgn2.6/ptodata/2/paa/US08_NEW_COMB.pcp.*
5: /cgn2.6/ptodata/2/paa/US09_NEW_COMB.pcp.*
6: /cgn2.6/ptodata/2/paa/US10_NEW_COMB.pcp.*
7: /cgn2.6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1693	48.2	904	1	PCT-US03-11231-18
2	666	19.0	854	5	US-09-350-841A-1589
3	117.5	3.3	704	5	US-09-674-546A-2655
4	114.5	3.3	704	5	US-09-674-546A-2653
5	105.5	3.0	3418	1	PCT-US03-24394-1
6	102	2.9	499	1	PCT-US03-28227-4455
7	101	2.9	1189	5	US-09-837-961A-6
8	101	2.9	1189	5	US-09-661-016B-8
9	100	2.8	590	6	US-10-425-114A-50697
10	100	2.8	870	7	US-60-490-890-130
11	100	2.8	1047	7	US-60-490-890-128
12	98.5	2.8	700	5	US-09-674-546A-2651
13	97.5	2.8	923	6	US-10-425-114A-57789
14	96.5	2.7	373	6	US-10-425-114A-49966
15	96.5	2.7	1176	5	US-09-837-961A-2
16	96.5	2.7	1934	1	PCT-US02-38594-136
17	96	2.7	1323	1	PCT-US03-19153-312
18	95.5	2.7	543	7	PCT-US03-24505-33
19	95.5	2.7	543	7	US-60-500-337-2014
20	95.5	2.7	543	7	US-60-500-337-2015
21	95.5	2.7	557	6	US-10-425-114A-45341
22	95.5	2.7	606	7	US-60-500-315-1024
23	95.5	2.7	690	5	US-09-908-576-49
24	95.5	2.7	690	5	US-09-904-553B-49
25	95.5	2.7	690	7	US-60-500-315-1025
26	95.5	2.7	691	7	US-60-500-315-1026

27	94.5	2.7	649	7	US-60-485-450-1430	Sequence 1430, Ap
28	94.5	2.7	649	7	US-60-485-450-1431	Sequence 1431, Ap
29	94.5	2.7	649	7	US-60-485-450-1432	Sequence 1432, Ap
30	94	2.7	574	6	US-10-425-114A-53512	Sequence 53512, A
31	94	2.7	577	6	US-10-425-114A-63109	Sequence 63109, A
32	94	2.7	621	6	US-10-425-114A-51628	Sequence 51628, A
33	94	2.7	1089	5	US-09-662-254B-18	Sequence 18, Appl
34	94	2.7	1130	1	PCT-US03-28227-5096	Sequence 5096, Ap
35	94	2.7	1144	1	PCT-US03-28227-3411	Sequence 3411, Ap
36	94	2.7	1179	1	PCT-US03-28227-3410	Sequence 3410, Ap
37	94	2.7	1224	1	PCT-US03-28227-5095	Sequence 5095, Ap
38	94	2.7	1231	1	PCT-US03-28227-3408	Sequence 3408, Ap
39	94	2.7	1233	1	PCT-US03-28227-5094	Sequence 5094, Ap
40	94	2.7	1234	1	PCT-US03-28227-5093	Sequence 5093, Ap
41	93.5	2.7	417	5	US-09-581-286A-412	Sequence 412, App
42	93.5	2.7	426	5	US-09-581-286A-288	Sequence 288, App
43	93	2.6	714	6	US-10-425-114A-50040	Sequence 50040, A
44	92.5	2.6	1259	5	US-09-976-858-238	Sequence 238, App
45	92.5	2.6	3037	7	US-60-500-315-1995	Sequence 1995, Ap

ALIGNMENTS

RESULT 1
PCT-US03-11231-18
; Sequence 18, Application PC/TUS0311231
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.53801PC
; CURRENT APPLICATION NUMBER: PCT/US03/11231
; CURRENT FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
PCT-US03-11231-18

Query Match		48.2%;	Score 1693;	DB 1;	Length 904;
Best Local Similarity		49.2%;	Pred. No. 4.3e-111;		
Matches 322;		Conservative 114;	Mismatches 194;	Indels 24;	Gaps 7;
QY	20	NSSPSTQNTVSREVSSVOLSEESTFYLCPPPVGTVIRLEPPKCPKPRKATENGEGI	79		
DB	82	NATVAAGHATLAHLREIKVENADAQFYVCPPTGATVYQEQPRCPTPEQONTTEGI	141		
QY	80	ALLFKNISPYKFKVLYYKNIIQTTWTGTYRQITNRYTDTPTYSIEITDLIDGKR	139		
DB	142	AVFKNIEAPYKATMYKIDVTVSQVWFGRHSQPMGFEDRAPVPFEEVDIKITKV	201		
QY	140	CSSKARYLNNTVVEAFDRDAGEKQVLLKPSFNTPESSAWHTTNTYTVWGSPTVYRTG	199		
DB	202	CRSTAKYVNNMTTAFHDDHETDMLPAKVATRTSGWHTTDLKYNPSRVEAFHRTG	261		
QY	200	TSVNCIVEEMDARSPEPSYFAMANGDIANSIPFYGLSPPEAAAPMGYPQDNFKOLDSY	259		
DB	262	TVNCLIVEEDARSVPYDEFLATGDFVMSPFYCYR-EGSHTSHTSYAADRKFQVDGF	320		
QY	260	FSDMLDKRKASLPVKARNFLITSHFTVGDWADPKTRVCSMTKWKKEVTEMLRATVNGRVR	319		
DB	321	YARDLTAKRATSPITRNLLITTPKFTVANDWVKRPVAVCTMTKWQVEWDLRAEYGGSR	380		
QY	320	FWARELSAFISNTTFEDPNRIILGOCIKREAAEAEIOIFRTKYNDSHVKGHVQYFLAL	379		
DB	381	FSSDAISTTTTNTLTETYSLSRDLGDCIGRDAREADRMFARKYNATHIKVGOPQPYLAT	440		

QY 449 TTAIKSTSSVQFAMQLQFLYDHIQTHINDMFSRIATAWCELQNLRELV----- 494
Db 526 LFGFHSFGNGDAVIDL-----TAGGETR-KELIRSGGSLSLNSING 567
QY 495 LMHEGKIN-----PS-----ATASATIGRRVAARKMLGDVAAVSS 529
Db 568 AWH-GIDMDNLKNGISGKTADNAAPSTPFHRETLNSELSDGISRHDITELFSDSLYVTS 626
QY 530 --CTAID---AESVTLQNSMR-----VITSTWTCYSRPLVLFYSYGENOQNIQQLG 575
Db 627 NGYTNLDTQELSDVLIRNAVHPKKNPIPLKITGT---VDKPSITVDYGRLTGGINSRKE 683
QY 576 ENNELPTL 584
Db 684 KOKILEDTL 692

RESULT 4
US-09-674-546A-2653
; Sequence 2653, Application US/09674546A
; GENERAL INFORMATION:
; APPLICANT: Institute for Genomic Research
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions
; FILE REFERENCE: CHIR-0334
; CURRENT APPLICATION NUMBER: US/09/674,546A
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 3264
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2653
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-674-546A-2653

Query Match 3.3%; Score 114.5; DB 5; Length 704;
Best Local Similarity 20.3%; Pred. No. 1.3;
Matches 136; Conservative 81; Mismatches 215; Indels 237; Gaps 33;

QY 87 ISPKFKVLYYKNI---IQTTWTGTYRQITNRYTDTPTVSIIEITDLIDGRCSS 142
Db 90 VSVQETKIGLSWKNLSWDQIQIEKVVSS---AELALTRDGKGVNQLDIDLSQKQAS 145
QY 143 KARYLRNNVYEAFFDRDAGEKQVLLKPSKEN--TPESRAWHHTNYTYWGS---PWIVR 197
Db 146 VNRIIVENSTVRL---NFLQEQILKEINLNQSPDSSGQPFESSGILVWGKLSVPWKR 202
QY 198 TGTSVNCIVEEMDARSVEPYSYFAMANGDIANI-----SPFYGLSPPEAAEPMGYPOD 251
Db 203 GLFLSNGI---GPPFISPFHFEASTSLDGHGITISTTSPSVYRFNAGGADAAGLGRAD 258
QY 252 -NFKQLDSYFMDLDRKRRKASLPVKRNFLTSHFTVGDWAP----- 292
Db 259 TSPRLHLTAQIPALALRNNSIKIE---TVNGAFTAGGEVARDWGSFKLDKANLHSGIAN 315
QY 293 -----KTR-----VCSTWKE-----VTEMLRAVYV--GRYFVAR-- 323
Db 316 IGNAEISGFKTPRHQTNFSLNSPLVWTKNGLDAPRLYVSTLQDVTNRLPQPRFISRLD 375
QY 324 -----ELSATISNT-----TEFDPNRIILGQCICKREAAEABOIFRTKY 363
Db 376 GSLSVPLNQNWAELNGTFDRQTVAAKFRYTHEDAPHL-----EAAVALQKLNLTYP 427
QY 364 ND-----SHVKVGHVO-----YFALGGFTVAYQPV 390
Db 428 LDDVROONGKIFPDTLAKLSGDIEAHKLGKGVQLPGLQDDMETYLHADKGIH-----AL 482
QY 391 SKSLAHMYRLMRDNRTEMLDLVNKHAIYKKNATSLRRLDRNRNAPNRKI---TLDD 448
Db 483 SRFKSGLY-----GGHTEGGISIANTRPATY-----RLQ---QNASNIQIQLLOD 525
QY 449 TTAIKSTSSVQFAMQLQFLYDHIQTHINDMFSRIATAWCELQNLRELV----- 494

Db 526 LFGFHSFGNGDAVIDL-----TAGGETR-KELIRSGGSLSLNSING 567
QY 495 LMHEGKIN-----PS-----ATASATIGRRVAARKMLGDVAAVSS 529
Db 568 AWH-GIDMDNLKNGISGKTADNAAPSTPFHRETLNSELSDGISRHDITELFSDSLYVTS 626
QY 530 --CTAID---AESVTLQNSMR-----VITSTWTCYSRPLVLFYSYGENOQNIQQLG 575
Db 627 NGYTNLDTQELSDVLIRNAVHPKKNPIPLKITGT---VDKPSITVDYGRLTGGINSRKE 683
QY 576 ENNELPTL 584
Db 684 KOKILEDTL 692

RESULT 5
PCT-US03-24394-1
; Sequence 1, Application PC/TUS0324394
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute
; APPLICANT: Ramin Shiekhatar
; TITLE OF INVENTION: METHODS FOR REGULATING BRCA1-BRCA2-CONTAINING COMPLEX ACTIVITY
; FILE REFERENCE: WSTR-0014B
; CURRENT APPLICATION NUMBER: PCT/US03/24394
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/401,433
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/449,950
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 3418
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-24394-1

Query Match 3.0%; Score 105.5; DB 1; Length 3418;
Best Local Similarity 17.7%; Pred. No. 32;
Matches 128; Conservative 130; Mismatches 247; Indels 219; Gaps 37;

QY 85 ENLSPYK--FKVLYYKNIQTTWTGTYRQITNRYTDTPTVSIIEITDLIDGRCSS 142
Db 1625 ENLTKSKIFLKVYKHENVEKETAKSPATC-----YTNOQSPYSVIENSALAFYTS--CSR 1677
QY 143 KA-----RYLRNNVYEAFFDRDAGEKQVLLKPSKFNTPESRAWHHTNETYTVWG 191
Db 1678 KTSVSTSLLEAKKWLREGI---FDG-----QPERINTAD-----YV 1711
QY 192 SPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMGYPOD 251
Db 1712 GNYLYENNS--NSTIAENDKNHLSERQDVTYLSNSSMSNSYSYHS---DEVYNDSGYLSK 1765
QY 252 NFKQLDSYFMDL---DKRRKASLPVKRNFLTSHFTVGDWAPKTR--VCSMTWKKE 305
Db 1766 N--KLDSGIEPVLKVNVEDQKNTSFSKVISNVKDANAI-----POTVNEDIC----- 1809
QY 306 VTEMLRATVNGRYRFMARELSATFISNTEFFDPN-----RIILQO---CIRKEAAAEQIF 359
Db 1810 VEELVTSSSPCKNKNAAIKLS---ISNSNFEVGPAPAFRIASGKIVCVSHETIKKVKDIF 1866
QY 360 RYKY-----NDSHVKVGHVQYFALGGFTVAYQPVLSKSL-----AHMYRLMRDNR 408
Db 1867 TDSFSKVIKENNENKSKICQTKIMAGCYEALDSEDIHLNSLNDDECSHSHKVFADIQS 1926
QY 409 DEMLDLVNKHAIYKKN-----ATSLSLRRLDRNRNAPNRKITLDDPTAK 453
Db 1927 EETLQHNQMSGLKRVSKISPCDVSLETSDICKCSIGKLHKSVSAA-----NTCGIF 1978
QY 454 STSS-----VQFAMQLQ---FLYDHIQTHINDMFSRIATAWCE-----LQNLREL 493
Db 1979 STASGKSVQVSDASLQNAQRFSEIEDSTKQVFSKVLFRKSNEHSQDLTRENTAINTPEH 2038

QY	494	VLWHGKIGK	---	INPSATA	--SATLGRVVA	-----AKMLGDVAASVCSCTAIDAESVT	539		
		: :	: :	: :	: :	: :	:		
Db	2039	LIISQKGSYVNVYSSAFSGFSTAGSKQVSTILESSLHKVGVLEEFDLIRTEHSLHYSPTS	2098						
QY	540	LQNSMRVITSTN	---	TCYSRPLVLFSYGENOGNIOQLGENN	---	ELLP	582		
		: :	: :	: :	: :	: :	:		
Db	2099	RQNVSKILPRVDRKRPNEHCNVSEMEKTC-SKEFKL	---	SNNLVGGSSNNHHSIKVSP	2153				
QY	583	TLFVPCPSANIRRRYFLFGSGVALFENYFNVKMDAA	---	DI	621				
Db	2154	YLSQFQ	---	QDKQQLVLGTQKVSLENHIVLHGREQASPKNVKMEIGKTETFSDPVKINI	2209				
QY	622	QIATSTF	---	VELNLTLEDEIRI	---	LPL	---	SVYKEELRDYGVLDYAE	661
		: :	: :	: :	: :	: :	: :	: :	:
Db	2210	EVGSTYSKDSNYFETEAVEIAKAFMEDDELDTSKLPFSHATHSLFTCPENEEM-VLSNSR	2268						
QY	662	VARR	665						
Db	2269	IGRR	2272						
		: :	:						
RESULT 6									
PCT-US03-28227-4455									
; Sequence 4455, Application PC/TUS0328227									
; GENERAL INFORMATION:									
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;									
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;									
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;									
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;									
; APPLICANT: ALTUS, Christa V.; PITTS, Steven J.;									
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;									
; APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;									
; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;									
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;									
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;									
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;									
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;									
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;									
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;									
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;									
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;									
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;									
; APPLICANT: XU, Yuming; KWONG, Mary;									
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;									
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;									
; APPLICANT: GITZEN, Daryl; PATURY, Srikanth.;									
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.;									
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS									
; FILE REFERENCE: PN-0100 PCT									
; CURRENT APPLICATION NUMBER: PCT/US03/28227									
; CURRENT FILING DATE: 2003-09-12									
; PRIOR APPLICATION NUMBER: US 60/410,260									
; PRIOR FILING DATE: 2002-09-12									
; PRIOR APPLICATION NUMBER: US 60/410,259									
; PRIOR FILING DATE: 2002-09-12									
; NUMBER OF SEQ ID NOS: 5444									
; SOFTWARE: PERL Program									
; SEQ ID NO 4455									
; LENGTH: 499									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: misc_feature									
; OTHER INFORMATION: Incyte ID No: 960393.PT29p									
PCT-US03-28227-4455									
Query Match 2.9%; Score 102; DB 1; Length 499;									
Best Local Similarity 19.5%; Pred. No. 6.5;									
Matches 109; Conservative 76; Mismatches 216; Indels 158; Gaps 27									
QY	163	KQVLKPKSKFNTP	---	SRAMHTTNTYTVWGSPWIRYTGTSVNCIVEMDARSVPYSY	219				
		: :	: :	: :	: :	: :	:		
Db	10	KSPATIKRNPTEPEPVSTGSHYGAEPTVSPCPSSSAKGTAVN	---	FSSLMTPTF	---	62			

RESULT 9
US-10-425-114A-50697
: Sequence 50697, Application US/10425114A
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: KOVALIC, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: US/10/425,114A
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 50697
: LENGTH: 590
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB3632-014-H9_Flt1.psp
US-10-425-114A-50697

Query Match 2.88: Score 100: DB 6: Length 590:

QY 375 YFLALGCF-IVAYOPVLKSLAHMYLRELMRDNRTDEMILVNNK--HAIYKKNATSLSR 431
Db 689 WFLSSHIPLKGIPEGLRVARYSMKIMPEEYSEFKELIQKELHVY-----ALSH 743
QY 432 LRRDIRNAPNRKIPLDDTTAKTSSTVQFAMQLQFLYDHIOHINDMESRIATAWCELQNR 491
Db 744 -----VCCQDRTILAS-----ILLRFLHEKLESLL-----ICTLNR 776
QY 492 ELVLWHEGIIKINPSATATLGRVAAKMLGDVAASCTAIDAEVTLQNS-MRVITST 550
Db 777 EISMEDEATTLFRATTLASTL-----MEQYMKATQATQPVHAKDSILKIMESK 825
QY 551 NTCYSRPLVLFESYGENOGTQQLGENNEL-----PTLEAVEPC---SAHR 595
Db 826 QSCELSPSKLEKNEDVNTLHLLNILSELVEKIFMASEILPPTLRYIYGCLOKSVQHK 884

RESULT 12
US-09-674-546A-2651
; Sequence 2651, Application US/09674546A
; GENERAL INFORMATION:
; APPLICANT: Institute for Genomic Research
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions
; FILE REFERENCE: CHIR-0334
; CURRENT APPLICATION NUMBER: US/09/674.546A
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 3264
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2651
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-674-546A-2651

Query Match 2.8% Score 98.5; DB 5; Length 700;
Best Local Similarity 18.68; Pred. No. 17;
Matches 128; Conservative 93; Mismatches 205; Indels 261; Gaps 34;

QY 76 GEGIAIFKENISPKFKVLYKNI-----IQTTWTGTTVYRQITNRYTDRTPVSIEIT 131
Db 85 GGRVAVSVKET-----KIGLSWKNLWSDRIQVEKVVVSGADLALTR--DRN--GANNIQ 134

QY 132 DLIDKGRCSKARYLRNNVVEAFDRDAGEKQVLLKPSKN--TPESRAWHHTNYTYV 189
Db 135 DLFDGAKHSASVNRRIIVENSTVRL--NFIQQQLILKEISLNQSPDSSGQFESSGILV 191

QY 190 W---GSPWIVRTGTSVNCIYEEMDARSVPYSYFAMANGDIANI-----SPFYGLSPPE 240
Db 192 WRKLSVPWKSRLGFLS---DGIGTPEISPFHFEASTSLDGHGIIITTGSPSVRFNAGG 247

QY 241 AAAPMGYPQDNFKQLDSYFSMDLDRKRSALPVKRNFL-----ITSHTVGDWAP--- 292
Db 248 ADAAGILGRADT-----SFRNLHTAQIPA-LALKNNSIKTGTVNGYTFAGGEVARDGS 301

QY 293 -----KTRVCSMTKWKVEMLRATVNG----- 316
Db 302 FKLDKANLHSGIANIGNAIEGSGFKPRL--QTNFSLGSLPWSRDNGLDAPRLHISTLQ 359

QY 317 -----RYRPMAR-----ELSAFF-----ISNTEFDPNRIILGQC 346
Db 360 DTVDRLPQPRFISRLDGSLSIPNLQNNWNAELNGTFDRQPAAKFKYTFEGAPH----- 412

QY 347 IKREAAAEIQIFRTKYND-----SHVKVGHVQY----- 375
Db 413 --LEAAALQKLNLAFLYDDEFROQNGKIFPDILGRLSGNVEAHLKIGSIQLPGQLDDME 470

QY 376 -----FLAL-----GGFIVA-YQPV---LSKSLAHMYLRELMRD----- 405
Db 471 TYLHADKDHIALSRFKSGLYGGHTEGGISIANTRPATYRLQQNASNIQIQLQDLQFGFH 530

QY 406 ---NRTDEMLDLV---NNKHAIYKKNATSLSRRLDRIRNAPNRKIPLDDTTAK- 453

Db 531 SFGSGDAVIDLITRASGENRKLIRSLQGSUS---LNISNGAWHGIDMD---SILKNGLSGK 585
QY 454 ---STSGVQFAMQLQFLYDHIOHIN--DMFSR-----IATAMCELQNRRLVLWHEGIIKINPS 505
Db 586 ISGSTPYRTLNSEISDGLSRHIDTELFSDSLVYVTSNGYTNLDTOEL----- 633
QY 506 ATASATLGRVAAKMLGDVAASCTAIDAEVTLQNSMR-----VIITSTNYCYSRP 557
Db 634 -----SEDLIRNAVHPKPKPIPLKITGT---VDKP 661
QY 558 LVLFESYGENOGTQQLGENNELPTL 584
Db 662 SITVDYGRLTGGLNSRKEKOKILEDITL 688

RESULT 13
US-10-425-114A-57789
; Sequence 57789, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57789
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC090C06_FLI.pep
US-10-425-114A-57789

Query Match 2.8% Score 97.5; DB 6; Length 923;
Best Local Similarity 20.38; Pred. No. 27;
Matches 116; Conservative 83; Mismatches 210; Indels 163; Gaps 28;

QY 175 PESRAWHT---TNETYTWGSPWYRTGTSYNCIYEEMDARSV----- 214
Db 200 PYDRIMSDLIKRNQYLVGVAPQGTERTISTTNRNIDTRETREYPPVKVMQTAVVGTGVLSYR 259

QY 215 -----FP-----YSYFAMANGDIANISFFYGLSPPEAA-----AEPMGYPQDNFKQ 255
Db 260 LNLEDFPANARAYAFABIEDLGQNESRKFPLKQPIYADYSNAYVNTAENANGSYTYLIE- 318

QY 256 LDSYFSMDLDRKRSALPVKRNFLITSHTVGDWAPKTR-----VCSMTKREV---T 307
Db 319 -PSYMNVTLE-----FVLSFSFVM---TPDSTRGPNLNALEISKYVQIASKT 361

QY 308 EMLRATVNGYRFMARELSATFISNTTEFPNRIILGQCIKRAE-----AAIEQI 358
Db 362 DKQDITTVNAFRLLSAQ-----SSQTNEGDP-----CVPTPWYVNCSTTTPPRITKI 409

QY 359 FRKYNDSHVKGHVQVFLALGGFIVAYQPVLSKLAHMYLRELMRDNRTDEMILVNNK 418
Db 410 ILSRN---YK-GEIPSELS-----NMALTELMDLGDLNLTGQLPDMSKLINLK 454

QY 419 HAIYKKNATSLSRRLDRIRNAPNRKIPLDDTTAKTSSTVQFAMQLQFLYDHIOHINDME 478
Db 455 -IVHLENKNTLGRPLSDMGSLPS-----LQALF---IQ---NSNF 487

QY 479 SRIATAWCELQNRRLV-----WHEGIIKINPSATATLGRVAAKMLGDVAASVSSCT 531
Db 488 SGEIPA--GLISKKIIFNYDGNAEHLRGGKKHFRWLGISIGVLVILLILFLVSLVLLN 545

QY 532 AIDAEVTLQNSMRVITSTNTCYSRPLVLFESYGENOGTQQLGENNELPTLEAVEPC 591

Db 546 TRRASKKREKIGSGRTN---SKPGYSFLRG---GNL-----MDENTTCYITLSELKEAT 596
 QY 592 ANHRRYFFGSGYALFENYFVKWDAADIOIAS-----TFVE--LNLTLLEDE 639
 Db 597 DNFSSKKIGKGS-----FGSVYTGKRDGKEIAVKSNESCHGNOQFVNEVALLSRHHRN 652
 QY 640 ILPLSVYTKELRDVGVDYAEVAR-RNOLHE 670
 Db 653 LVPLIGYCEECQHILVYEHMNGTLRDIHE 684

RESULT 14
 US-10-425-114A-49966
 ; Sequence 49966, Application US/10425114A
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114A
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 49966
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700239581_FLI.pap
 US-10-425-114A-49966

Query Match 2.7%; Score 96.5; DB 6; Length 373;
 Best Local Similarity 18.9%; Pred. No. 11;
 Matches 83; Conservative 60; Mismatches 162; Indels 135; Gaps 16;

QY 29 TSREVVSSVOLSEESTFYLCPPPGVSTVIRLEP-----PRKCEPRKATEW 75
 Db 3 SORTSLVVKMGVRSYFFIIAANVAVVALLPAGATGAGLVGYFYSKTCP---SAESL 59
 QY 76 GEGIALFKENISPYKFKVLYKNIQTTWTGTYRQITNRYTDRTPVSTEEITDLID 135
 Db 60 VQAVAAAFKNSGIAAGLRLHFHDCFV-----RGCDGSV-----LTD 97
 QY 136 GKGRCSKARYLRNNVYVEAFD-RDAGEKOVLLKPSK-----FNPESRAWHHTNE 185
 Db 98 STANTAEKDAVPPNPSLRGFEVIDAAKRAVEARCPKTVSCADILAFARDSIALAGNML 157
 QY 186 TTVVMSPIWYRTGTSVNCIVEEMDARSFYPSYFAMANGDIANTISPFYGLSPPEAAAP 245
 Db 158 TYKV---PAGRRDR---VSRDTDANSNP-----SPLSTAEL 190
 QY 246 MGYPDQNFQLODYSFMSDLDRKASLPVKNFLITSHFTVGDWAPK-TTRVCSMTWK 304
 Db 191 VG---NF-----TRKNLTAEDEVVLSGAHTVGRSHCSSTNRLYGFSNAS 232
 QY 305 EVTEMLRATVNGRYRPMARELSATFISNTTEFDPN----- 339
 Db 233 DVD-----PIISSAYALLRAICP---SNTSQFPNTTMDLITPALLDNRYVGLANNL 285
 QY 340 -----RIILGQCIRKREAAAI--EQIFRTKYNDSHVKVGVQVYFALGGFIVAYQPV 390
 Db 286 GLFTSDQALLTNATLKSDVAFVKSSAWTKFAKSMVKMGINDVLTGTKEIRLCNRCVI 345
 QY 391 SKSLAHMYLRMLRDNRTDE 410
 Db 346 NSGSSSSGLFQLHTATASDE 365

RESULT 15
 US-09-837-961A-2
 ; Sequence 2, Application US/09837961A
 ; GENERAL INFORMATION:
 ; APPLICANT: Payne, Jewel
 ; APPLICANT: Sick, August J.
 ; TITLE OF INVENTION: Novel Bacillus thuringiensis Isolate Active Against Lepidopter
 ; TITLE OF INVENTION: Pests, and Genes Encoding Novel Lepidopteran-Active Toxins
 ; FILE REFERENCE: MA-43CDF2D3
 ; CURRENT APPLICATION NUMBER: US/09/837,961A
 ; CURRENT FILING DATE: 2001-04-19
 ; PRIOR APPLICATION NUMBER: US 09/521,344
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: US 08/933,891
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: US 08/356,034
 ; PRIOR FILING DATE: 1994-12-14
 ; PRIOR APPLICATION NUMBER: US 08/210,110
 ; PRIOR FILING DATE: 1994-03-17
 ; PRIOR APPLICATION NUMBER: US 07/865,168
 ; PRIOR FILING DATE: 1992-04-09
 ; PRIOR APPLICATION NUMBER: US 07/451,261
 ; PRIOR FILING DATE: 1989-12-14
 ; PRIOR APPLICATION NUMBER: US 371,955
 ; PRIOR FILING DATE: 1989-06-27
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 1176
 ; TYPE: PRT
 ; ORGANISM: Bacillus thuringiensis
 US-09-837-961A-2

Query Match 2.7%; Score 96.5; DB 5; Length 1176;
 Best Local Similarity 18.6%; Pred. No. 41;
 Matches 118; Conservative 94; Mismatches 213; Indels 211; Gaps 36;

QY 120 TDRTPVSTEEITDLI-----DGKGRCSKARYLRNNVYVEAFDRDAGEKOVLLKPS 170
 Db 634 TNVTDYHDQVSNLVECLSGFELDEKRELSEKVAHA-----NLSDERNLLQDPN 684
 QY 171 --KFTNPESRAWHHTNRYTVWGSPIYR-----TGTSVNC-----IYEMDARSVPFYS 218
 Db 685 PRGINRQPDGRWGSTD-ITIOGGDDVFKENYVTLPGTFNECYPYLYKQIDESKLYAT 743
 QY 219 YFAMANGDIANTISPFYGLSPPEAAAPMGYQODNFQKLDYSFMSDLDRRKRASLPVKRNF 278
 Db 744 RYQL-----RGYIEDS-QHLEIYL-----IRYN- 765
 QY 279 LITSHFTVGDWAPKTRVCSMTKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFD 338
 Db 766 --TKHETVN---VPGTGLMPLSVENPIGKC-----GEPNRCAPQL-----ENP 805
 QY 339 NRILGQCIRKREAAAEIQ-----IFRTKYNDSHVKVGVQV 375
 Db 806 DL-----DCSRDGEKCAHSHHFSLDIDTIGCTDLNENLGVWVIFIKKQDGHARLGNLEF 861
 QY 376 FLALGGFIVAYQPVLSKSLAHMYLRE-LMRDNRDTEMLDVLNKHAIYKKNATSL----- 429
 Db 862 -----LEEAPLVGESLARVKRAEKWRDRR--EKLOVETN--IVYKEAKESVDALFV 909
 QY 430 -----SRLRRD-----IRNAPN-----RKITLDDTTAISTSSVQFAMLQ-----FLYD 468
 Db 910 NSQYDLQADTDIAMTHAADKRVHIREAYLPFELSVIPGVNAGIFEELGRIFFAYSLYD 969
 QY 469 HTQTHINDMFSRIATAW-----CELQNRRELVLHHEGKIKNPSATASATGLRRVAAKML 521
 Db 970 ARNVKNGDFNGLSCWNVKGHVDEQNN-----HRSVLVYPEWEEAEVSEQVRVCPGR- 1023
 QY 522 GDVAAVSS-----CTAI-----DAESVTIQNSMR---VITSTNTCYSRPLVLFSGY- 564
 Db 1024 GYLIRVTAYKEGYGCGCVTIHEIDNTDELKFSNCVEEVEEVPNTVTNCTNDYANOEEYGG 1083

QY 565 -----ENQNIQOLGENNELLPTLEA-----VEPCAN--HRRYFLGSGYAL 606
 Db 1084 AYTSRNRG--YGESYESNSSIPAAYAPVYEEAYIDGRKENPCESNRGYGDTPLPAGYVT 1141
 QY 607 FENYNEVKMWDAAIQIA---STFV--ELNLTILED 637
 Db 1142 KE-LEYFPETDKRWIEIGETEGTFIVDSVELLMEE 1176

Search completed: October 8, 2003, 17:25:53
 Job time : 16.6904 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:00 ; Search time 269.154 Seconds
(without alignments)
2271.804 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672
Perfect score: 3510
Sequence: 1 MHFRNCIFFLVILYGTN.....DVGVDYAEVARRNQLHEK 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_Main.*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	3510	100.0	865 3	US-07-722-860-13 Sequence 13, Appl

2	3510	100.0	865 6	US-08-213-449A-13 Sequence 13, Appl
3	3510	100.0	865 6	US-08-213-449B-13 Sequence 13, Appl
4	3510	100.0	1086 15	US-09-147-052-4 Sequence 4, Appl
5	1726	49.2	868 12	US-08-804-439-21 Sequence 21, Appl
6	1726	49.2	868 17	US-09-301-390-21 Sequence 21, Appl
7	1726	49.2	868 17	US-09-338-326-21 Sequence 21, Appl
8	1699	48.4	891 9	US-08-541-878-6 Sequence 6, Appl
9	1698.5	48.4	933 3	US-07-805-524-2 Sequence 2, Appl
10	1693	48.2	904 1	PCT-US03-11231-18 Sequence 18, Appl
11	1693	48.2	904 27	US-10-121-988-18 Sequence 18, Appl
12	1693	48.2	904 28	US-10-200-562-18 Sequence 18, Appl
13	1693	48.2	904 28	US-10-237-551-18 Sequence 4, Appl
14	1683.5	48.0	943 4	US-08-096-183D-4 Sequence 8, Appl
15	1683.5	48.0	943 19	US-09-521-738-14 Sequence 23, Appl
16	1679	47.8	885 9	US-08-541-878-8 Sequence 23, Appl
17	1674	47.7	885 12	US-08-804-439-23 Sequence 22, Appl
18	1674	47.7	885 17	US-09-301-390-23 Sequence 22, Appl
19	1674	47.7	885 17	US-09-338-326-23 Sequence 22, Appl
20	1660	47.3	903 12	US-08-804-439-22 Sequence 22, Appl
21	1660	47.3	903 17	US-09-301-390-22 Sequence 22, Appl
22	1660	47.3	903 17	US-09-338-326-22 Sequence 12, Appl
23	1654	47.1	904 32	US-60-412-956-12 Sequence 142, App
24	1652	47.1	846 5	US-08-123-456-142 Sequence 142, App
25	1652	47.1	846 16	US-09-297-477A-142 Sequence 142, App
26	1652	47.1	846 25	US-09-394-404-142 Sequence 241, App
27	1652	47.1	896 5	US-08-123-456-241 Sequence 241, App
28	1652	47.1	896 16	US-09-297-477A-241 Sequence 241, App
29	1652	47.1	896 25	US-09-394-404-241 Sequence 79, Appl
30	1623	46.2	787 5	US-08-123-456-79 Sequence 79, Appl
31	1623	46.2	787 16	US-09-297-477A-79 Sequence 79, Appl
32	1623	46.2	787 25	US-09-394-404-79 Sequence 2, Appl
33	761	21.7	907 4	US-08-048-978-2 Sequence 19, Appl
34	761	21.7	907 12	US-08-804-439-19 Sequence 19, Appl
35	761	21.7	907 17	US-09-301-390-19 Sequence 19, Appl
36	761	21.7	907 17	US-09-338-326-19 Sequence 2, Appl
37	761	21.7	907 28	US-10-233-538-2 Sequence 2, Appl
38	745	21.2	830 26	US-10-055-364-45 Sequence 24, Appl
39	743	21.2	865 1	PCT-US00-18647-24 Sequence 13, Appl
40	743	21.2	865 26	US-10-055-364-24 Sequence 13, Appl
41	738.5	21.0	713 32	US-60-435-549-13 Sequence 11, Appl
42	738.5	21.0	906 32	US-60-435-549-11 Sequence 55, Appl
43	737.5	21.0	793 26	US-10-055-364-55 Sequence 40, Appl
44	733	20.9	844 26	US-10-055-364-40 Sequence 15, Appl
45	733	20.9	874 12	US-08-804-439-15

ALIGNMENTS

RESULT 1
US-07-722-860-13
; Sequence 13, Application US/07722860
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOXP2 VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 North Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/722,860
: FILING DATE: 19910628
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1644-101P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 241-1300
: TELEFAX: (703) 241-2848
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 865 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-722-860-13

Query Match 100.0%; Score 3510; DB 3; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSVSVQSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSVSVQSEESTFYLCPPVGVSTVIRL 60
QY 61 EPRKCPKPRKATEWEGEGTALFKENISPKYKVTLYYKNIITQTTWTGTTTQITNRYT 120
DB 61 EPRKCPKPRKATEWEGEGTALFKENISPKYKVTLYYKNIITQTTWTGTTTQITNRYT 120
QY 121 DRTPVSEETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKENTPESRAW 180
DB 121 DRTPVSEETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKENTPESRAW 180
QY 181 HTTNETYTVWGSPIWYRTGTSVNCIYEEMDARSVPFYSFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVWGSPIWYRTGTSVNCIYEEMDARSVPFYSFAMANGDIANISPFYGLSPPE 240
QY 241 AAAPMGYPQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGVWMAKPTTRVCSM 300
DB 241 AAAPMGYPQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGVWMAKPTTRVCSM 300
QY 301 TKWKEVTEMLRAVNGRYRFMARELSATFISNTTEFPDNRILIGQCILKREAAAEIQIFR 360
DB 301 TKWKEVTEMLRAVNGRYRFMARELSATFISNTTEFPDNRILIGQCILKREAAAEIQIFR 360
QY 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDRNTDMLDLVNNKHA 420
DB 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDRNTDMLDLVNNKHA 420
QY 421 IYKKNATSLRLDRIRNAPNRKITLDDTTAISTSSVQFAMQLQFLYDHIQTHINDMFSR 480
DB 421 IYKKNATSLRLDRIRNAPNRKITLDDTTAISTSSVQFAMQLQFLYDHIQTHINDMFSR 480
QY 481 IATAWCELQNLRELVLWHEGINKINPSATSLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
DB 481 IATAWCELQNLRELVLWHEGINKINPSATSLGRVAAKMLGDVAAVSSCTAIDAESVTL 540
QY 541 QNSMRVITSTWTCYSRPLVLFSGYENQGNIGQIGENNELPTLEAVEPCSANHRRYFLF 600
DB 541 QNSMRVITSTWTCYSRPLVLFSGYENQGNIGQIGENNELPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKMVDAAADIQIASTFVELNLTLLEDRILPLSVYKKEELRDVGVLIDYA 660
DB 601 GSGYALFENYFVKMVDAAADIQIASTFVELNLTLLEDRILPLSVYKKEELRDVGVLIDYA 660
QY 661 EVARRNQLHELK 672
DB 661 EVARRNQLHELK 672

RESULT 2
US-08-213-449A-13
: Sequence 13, Application US/08213449A
: GENERAL INFORMATION:
: APPLICANT: NAZERIAN, Keyvan
: APPLICANT: LEE, Lucy F.
: APPLICANT: YANAGIDA, Noboru
: APPLICANT: OGAWA, Ryohei
: APPLICANT: LI, Yi
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
: PROTECTION AGAINST MAREK'S DISEASE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/213,449A
: FILING DATE: 15-MAR-1994
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1644-108P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 865 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-213-449A-13

Query Match 100.0%; Score 3510; DB 6; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSVSVQSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSVSVQSEESTFYLCPPVGVSTVIRL 60
QY 61 EPRKCPKPRKATEWEGEGTALFKENISPKYKVTLYYKNIITQTTWTGTTTQITNRYT 120
DB 61 EPRKCPKPRKATEWEGEGTALFKENISPKYKVTLYYKNIITQTTWTGTTTQITNRYT 120
QY 121 DRTPVSEETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKENTPESRAW 180
DB 121 DRTPVSEETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKENTPESRAW 180
QY 181 HTTNETYTVWGSPIWYRTGTSVNCIYEEMDARSVPFYSFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVWGSPIWYRTGTSVNCIYEEMDARSVPFYSFAMANGDIANISPFYGLSPPE 240
QY 241 AAAPMGYPQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGVWMAKPTTRVCSM 300
DB 241 AAAPMGYPQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGVWMAKPTTRVCSM 300
QY 301 TKWKEVTEMLRAVNGRYRFMARELSATFISNTTEFPDNRILIGQCILKREAAAEIQIFR 360
DB 301 TKWKEVTEMLRAVNGRYRFMARELSATFISNTTEFPDNRILIGQCILKREAAAEIQIFR 360
QY 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDRNTDMLDLVNNKHA 420
```


Db 361 TKYNDSHVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDVLNNKHA 420
 QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 QY 481 IATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDAVAVSSCTAIDAESVTL 540
 Db 481 IATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDAVAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITNTCYSRPLVLSYGENQNIQGGQGENNELLPTLEAVEPCSAHRRYFLF 600
 Db 541 QNSMRVITNTCYSRPLVLSYGENQNIQGGQGENNELLPTLEAVEPCSAHRRYFLF 600
 QY 601 GSGYALFENYFVKVMDAADIQIASTFVLELNLTLLEDEILPLSVYTKELRDVGVLDA 660
 Db 601 GSGYALFENYFVKVMDAADIQIASTFVLELNLTLLEDEILPLSVYTKELRDVGVLDA 660
 QY 661 EVARRNQLHELK 672
 Db 661 EVARRNQLHELK 672

RESULT 3

US-08-213-449B-13
 ; Sequence 13 Application US/08213449B
 ; GENERAL INFORMATION:
 ; APPLICANT: NAZERIAN, Keyvan
 ; APPLICANT: LEE, Lucy F.
 ; APPLICANT: YANAGIDA, Noboru
 ; APPLICANT: OGAWA, Ryohel
 ; APPLICANT: LI, Yi
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
 ; PROTECTION AGAINST MAREK'S DISEASE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/213,449B
 ; FILING DATE: 15-MAR-1994
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M.
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 1644-108P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 865 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-213-449B-13

Query Match 100.0%; Score 3510; DB 6; Length 865;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILYIGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGTVIRL 60
 Db 1 MHYFRNCIFFLVILYIGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGTVIRL 60

QY 61 EPPKCPERKATEWEGEGIAILFKENISPYKFKVTLIYKNIQTTTWTGTYQIINRYT 120
 Db 61 EPPKCPERKATEWEGEGIAILFKENISPYKFKVTLIYKNIQTTTWTGTYQIINRYT 120
 QY 121 DRTPVSIETDLDIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPPESRAW 180
 Db 121 DRTPVSIETDLDIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPPESRAW 180
 QY 181 HTTNETYTVMGSPWIIYRTGTSVNCIIVEEMDARSVFYPSYFAMANGDIANTISPYGLSPPE 240
 Db 181 HTTNETYTVMGSPWIIYRTGTSVNCIIVEEMDARSVFYPSYFAMANGDIANTISPYGLSPPE 240
 QY 241 AAEPNGYQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGMWNAKTRTRVCSM 300
 Db 241 AAEPNGYQDNFKQLDSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGMWNAKTRTRVCSM 300
 QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDPNRIILGQCICKREAAIEQIFR 360
 Db 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFPDPNRIILGQCICKREAAIEQIFR 360
 QY 361 TKYNDSHVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDVLNNKHA 420
 Db 361 TKYNDSHVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDVLNNKHA 420
 QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 QY 481 IATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDAVAVSSCTAIDAESVTL 540
 Db 481 IATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDAVAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITNTCYSRPLVLSYGENQNIQGGQGENNELLPTLEAVEPCSAHRRYFLF 600
 Db 541 QNSMRVITNTCYSRPLVLSYGENQNIQGGQGENNELLPTLEAVEPCSAHRRYFLF 600
 QY 601 GSGYALFENYFVKVMDAADIQIASTFVLELNLTLLEDEILPLSVYTKELRDVGVLDA 660
 Db 601 GSGYALFENYFVKVMDAADIQIASTFVLELNLTLLEDEILPLSVYTKELRDVGVLDA 660
 QY 661 EVARRNQLHELK 672
 Db 661 EVARRNQLHELK 672

RESULT 4

US-09-147-052-4
 ; Sequence 4 Application US/09147052
 ; GENERAL INFORMATION:
 ; APPLICANT: SAITOH, Shuji
 ; APPLICANT: TSUZAKI, Yoshinari
 ; APPLICANT: YANAGIDA, Noboru
 ; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
 ; RECOMBINANT VIRUS, AND ITS USE
 ; FILE REFERENCE: 981167
 ; CURRENT APPLICATION NUMBER: US/09/147,052
 ; CURRENT FILING DATE: 1999-04-05
 ; PRIOR APPLICATION NUMBER: JP 08-103548
 ; PRIOR FILING DATE: 1996-03-29
 ; PRIOR APPLICATION NUMBER: PCT/JP97/01084
 ; PRIOR FILING DATE: 1997-03-28
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1086
 ; TYPE: PRT
 ; ORGANISM: hybrid
 ; US-09-147-052-4

Query Match 100.0%; Score 3510; DB 15; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MHYFRRNCIFELVILYGTNSSTQNTSVTREVSSVQLSEESTFYLCPPPVGSTVIRL 60
DB 1 MHYFRRNCIFELVILYGTNSSTQNTSVTREVSSVQLSEESTFYLCPPPVGSTVIRL 60
QY 61 EPPKPCPEPRKATWEGGIAILFKENISPKYKFKVTLTYKNIQTTWTGTYRQITNRYT 120
DB 61 EPPKPCPEPRKATWEGGIAILFKENISPKYKFKVTLTYKNIQTTWTGTYRQITNRYT 120
QY 121 DRTPVSTEEITDLDGKRGSSKARYLRNNVYEFADRDAGEKOVLLKPSKFNTPESRAW 180
DB 121 DRTPVSTEEITDLDGKRGSSKARYLRNNVYEFADRDAGEKOVLLKPSKFNTPESRAW 180
QY 181 HTTNETTYVMSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPYGLSPPE 240
DB 181 HTTNETTYVMSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPYGLSPPE 240
QY 241 AAAPPMGYPQDNFQKQDSYFMDLDRKASLPVKRNFLITSHFTVGDWNAKPTRVCSM 300
DB 241 AAAPPMGYPQDNFQKQDSYFMDLDRKASLPVKRNFLITSHFTVGDWNAKPTRVCSM 300
QY 301 TKWEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360
DB 301 TKWEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360
QY 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKHA 420
DB 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKHA 420
QY 421 IYKKNATSLRLDRIRNAPNRKITLDDTTAKSTSSVQFAMLOFLYDHIQTHINDMESR 480
DB 421 IYKKNATSLRLDRIRNAPNRKITLDDTTAKSTSSVQFAMLOFLYDHIQTHINDMESR 480
QY 481 IATAWCELQNLRELVLWHEGIIKINFSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
DB 481 IATAWCELQNLRELVLWHEGIIKINFSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
QY 541 QNSMRVITSTWTCYSRPLVLSYGENOQIOLGKNNELLPTLEAVEPCSANHRRYFLF 600
DB 541 QNSMRVITSTWTCYSRPLVLSYGENOQIOLGKNNELLPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKMVDAAQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVDYA 660
DB 601 GSGYALFENYFVKMVDAAQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVDYA 660
QY 661 EVARNQNLHELK 672
DB 661 EVARNQNLHELK 672

```

```

RESULT 5
US-08-804-439-21
; Sequence 21, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-21

```

```

Query Match 49.2%; Score 1726; DB 12; Length 868;
Best Local Similarity 50.5%; Pred. No. 4.2e-167;
Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;

QY 31 REVSSVQLSEESTFYLCPPPVGSTVIRLEPPKPCPEPRKATWEGGIAILFKENISPY 90
DB 41 REATHKSQDAETKFTFYVCPPTGTIVRLPTCTCPHLGKNFTGEGIAVYKENTAA 100
QY 91 KFKVTLTYKNIQTTWTGTYRQITNRYTDRTPVSTEEITDLDGKRGSSKARYLRNN 150
DB 101 KFKATVYKDVIVTAWAGSSYQITNRYADRVPIPVSEITDITDKFKCCKSSKATVYRN 160
QY 151 VYVAFDRDAGEKOVLLKPSKFNTPESRAHITNETTYVMSPIWYRTGTSVNCIVEEMD 210
DB 161 HKVEAFNEKDPQMDPLIASKYNVSGKAWHTTNDTYNAGTGTGTGTGTGTGTGTGT 220
QY 211 ARSVPYPSYFAMANGDIANISPPYGLSPPEAAAPPMGYPQDNFQKQDSYFMDLDRKKA 270
DB 221 ARSIPFDVSGLSGDIYMSPPFGLR-DGAYRHSNAYAMDRFHQFEGYRQRLDTRALL 279
QY 271 SLPVKRNFLITSHFTVGDWNAKPTRVCSMTKWEVTEMLRATVNGRYRPMARELSATFI 330
DB 280 E-PAARNFLVTPHLLTVGNWPKRTEVCSLVKREVEDVVRDEYAHNFRPTMKTLSTTFI 338
QY 331 SNTTEFDPNRIILGQCICKREAAAEIQIFRKYNDSHVKVGHVQYFLALGGFIVAYQPV 390
DB 339 SETNEFNQIHLSDQVKEARAIINRIYTRYNSSHVRTGDIQIYTLARGGFVVFQPLL 398
QY 391 SKSLAHMYLRELMRDNRDTEMLDLVNNKHAIVKKNATSLRLDRIRNAPNRKITLDDTT 450
DB 399 SNSLARLYQLQELVRE-----NTNHS-PQKHTPTNRSRSV-----PVELRANR 441
QY 451 AIKSTSSVQFAMLOFLYDHIQTHINDMESRIATWCELQNLRELVLWHEGIIKINFSATASA 510
DB 442 TITTTSSVEFAMLOFTYDHIQEHVNEMLARISSSQCLQNRERALLWSGLFPIINPSALAST 501
QY 511 TLGRVAAKMLGDVAASVSCTAIDAES-VTLQNSMRVITSTWTCYSRPLVLSYGENOQ 569
DB 502 ILDQVRKARILGDVIVSNCPGLSGDTRIILQNSMRVSTGTRCYSRPLISTVLSNGSGT 561
QY 570 IQGQLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVDAAQIASTFVE 629
DB 562 VEGQLGTQNLINMSRDLLEPCVANHKRYFLFGHHVYEDYRYVREIAVHDVGMLSTYVD 621
QY 630 LNLTLLEDEILPLSVYTKKEELRDVGVDYAEVARNQNLHELK 672
DB 622 LNLTLKDRFMPLOVYTRDELRTGLDYSIQRRNQMSLR 664

```

```

RESULT 6
US-09-301-390-21
; Sequence 21, Application US/09301390
; GENERAL INFORMATION:

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	49.2%	Score 1726;	DB 17;	Length 868;
Best Local Similarity	50.5%	Pred. No. 4.2e-167;		
Matches 325;	Conservative 116;	Mismatches 182;	Indels 20	

Qy	31	REVSVSQLSEEBSTFYLCPPPPGSGTWRLEPPRKCPPEPKRATBGEIGAILFKENISPY	90
Db	41	REALHKSDAETKPTFFVCPPTGSTIVRLEPTCTCPDYHLGKNFTEGIAVVYKENIAY	100
Qy	91	KFKVTLYYKNIQTTWTGTYTQITNRYTDRTPVSEETETDLIDGKRCRCSKARYLNN	150
Db	101	KFKATVYKQIVSTAWAGSYTOITNRYADRVIPVSEITDTDKFKGCSKATYVNN	160
Qy	151	VYVEAFORDAGEKOVLLKPSKFNPESSRAHHTTNETITVWGSPTIYRTGTSVNCIVERMD	210
Db	161	HKVEAFNEDNQPDMPLIASKYNSVSGKAHHTTNDTYAVAGTPTGYRTGTSVNCILIEVE	220
Qy	211	ARSVPYSYFAMANGDIANTISPYGLSPPEAAEPMGYPDQNFQKLOSYFSMDLDRKKA	270
Db	221	ARSIFPYDSFGLSTGDIYIYSPFGLR-DGAYREHSNYAMDRLFHQFEGYQRQDLDRALL	279
Qy	271	SLPKRNFLTSHPTGVCWDWAPKTRVCSMTKWEVTEMLRATVNGRVRFWARELSATFI	330
Db	280	E-PAARNFLVPHLTUVGNWKNPKRTECSLVKWEVEVDVDEYAHNRF*TMKLTSTTFI	338
Qy	331	SNTEFOPNRITILQCCKREAAAEIQFTKRYNDSHVKVGHVGYFALGCGFIYAYQPVL	390
Db	339	SETNEFNLIQHLSQCVCKEEARAIINRIYTRYNSSHVRTGDIQTYLARGGFVVVFQPLL	398
Qy	391	SKSLAHMYLRELMRNDTDEMLDVUNKNKHATYKKNATSLSLRLDRDIRNAPNRKITLDOTT	450
Db	399	SNLSRLYILQELVRE-----NTNHS-PQKHPTNTRSSRSV-----PVELRANR	441

```

451 AIKSTSSVQFAMQLFDYDHIOTHINDMFSRATANCELQNRELVLWHEGKINPSATASA 510
442 TITTTSSVEFAMQFTYDHIQEHVNEMLARISSWCQLQNRERALWSGLFPINPSALAST 501
511 TIGRRVAKMLGDVAAVSSCTAIDAES-VTLQNSMRVITSTWTCYSRPLVLSYGENQGN 569
502 ILDQVRKARIUGDVISVNCPELGSDDTRIILQNSMRVSGSTTRCYSRPLSTIVSLNGSGT 561
570 IQQOLGENNELPTLEAVEPCSANRRRYFLFGSGYALFENYFVKMVDAAADIQIASTFEVE 629
562 VEGQLGTDNELMSRDLLPEPCVANHKRYFLFGHHVYVEDYRVREIAVHDVGMISTYVD 621
530 LNLILLEDEILPLSVYTKRELRDVGVDYAEVARRNQLHEUK 672
622 LNLTLKDRFEMPLQVYTRDELRTGLLDYSEIQRNQMHSLR 664

RESULT 7
US-09-338-326-21
; Sequence 21, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-326-21

```

	Query Match	49.2%	Score 1726;	DB 17;	Length 868;
	Best Local Similarity	50.5%;	Pred. No. 4.2e-167;		
	Matches 325;	Conservative 116;	Mismatches 182;	Indels 20;	Gaps 6
Qy	31	REVSSVQLSSEESTFYLCPPPVGVSTVRLPPRKCPEPRKATWEGEGIALFKENISPY	90		
		: : : :			
Db	41	REALKSDAETKPTFYCCPPPTGVTVRLPPTCTCDPYHLGKNFTGPIAVVYKNIAY	100		
Qy	91	KFKVTLYYKNIQTWTGTVTRQITNNRYTORTPVYSIEETDLIDGKRCSSKARYLANN	150		

Query Match 49.28; Score 1726; DB 17; Length 868;

Query Match	49.2%	Score 1726;	DB 177;
Best Local Similarity	50.5%	Pred. No. 4.2e-167;	

Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;

QY	31	REVSSVQLSEBESTFYLCPPVGVSTVIRLEPPRCPEPRKATWEGEGIALLFKENISPY	90
Db	41	REALTKSQDAETKTFYVCPPTGTVIRLEPTRCTPDYHLGKNFTGEGIAVYVKENIAAY	100

91 KFKVTLYYKNIIQTTTWGTYYRQITNRYTDRTPVSEIETDLIDGKGRCSSKARYLRNN 150

[illegible]

101 KFKATVYKDVIVSTAWAGSYQTITNRYADRPVPIVSEITDIDKFGKSSKATYVRNN 160
151 VYVEAFDRDAGEKOVLLKPKSFNPESRAWHITNETYVVGSPWYRTGTSGVNCIVEEMD 210
161 HKVEAFNEDKNPQDPLIASKYNSVSGKANHTNDTYVAGTCTYTGTSVNCIIIEVE 220
211 ARSVEPYYSFAMANGDIANTISPFYGLSPPEAAAPMGYPQDNFKQLDSYFMSLDKRRKA 270
221 ARSIFPYDSFGLSTGDIYNSPPFGLR-DGAYREHSNVAAMDREHOFEGYQRDLDRALL 279
271 SLPVKRNPLTISHTVGDWNAKPKTRVCSTKWKVEVTEMLRATVNGRYRPMARELSATFI 330
280 E-PAARNEFLVTPHITVGNWPKPKTEVCSLVKWEVEDVVRDEYAHNFRFTMKLTSTFI 338
331 SNTTEFPDNRITLIGOCIKREAAAEIOIFRTKYNDSHVKGHGVYFLALGGFIVAYOPVL 390
339 SETNEFLNLQHLSCQVKEARALINRIYTRYNSHVRGDIQTYLARGGFVVFPQLL 398
391 SKSLAHMYLRELMRDNRTDEMLDLVNNKHAIYKKNATSLSLRDRINAPNRKITLDDTT 450
399 SNSLARLYQLQELVRE-----NTNHS-PQKHPTNTRSRSV-----PVELRANR 441
451 AIKSTSSVQFAMLOFLYDHIQTHINDMFESRTATWCELQNLRELVLWHEGINKINPSATASA 510
442 TITTTSSVEFAMLOFTYDHIQTHINDMFESRTATWCELQNLRELWGLFPIFNSALAST 501
511 TLGRVRAAKMGDVAASVSSCTAIDAES-VTLQNSMRVITSTNTCYSLRPLVLFYSGENQGN 569
502 ILDRQVAKARILGDVISVNCPELGSDRITLQNSMRVSGSTTRCYSLRPLISIVLSNGST 561
570 IQQGLGENNELLPTELEAVPCSAHRRYFLGSGYALFENYFVKWDAADIIQIATSTVE 629
562 VEGQGLTGNELTMSDLELPCVANHKRYFLFGHHYVYEDRYVREIAVHDVGMISTYVD 621
630 LNLTLLEDEILPLSVYKKEELRDVGVDYAEVARRNQLHELK 672
622 LNLTLKDRFMPQLQVYTRDELRTGLDLYSEIQRNQMSLR 664

RESULT 8
US-08-541-878-6
; Sequence 6, Application US/08541878
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Black, Darla
; APPLICANT: Scinicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cox & Smith Incorporated
; STREET: 112 East Pecan Street, Suite 2000
; CITY: San Antonio
; STATE: Texas
; COUNTRY: USA
; ZIP: 78205
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,878
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 08/042,747
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haymond, W. Bradley
; REGISTRATION NUMBER: 35186

REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 891 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-541-878-6

Query Match 48.4%; Score 1699; DB 9; Length 891;
Best Local Similarity 49.2%; Pred. No. 2.7e-164;
Matches 322; Conservative 99; Mismatches 210; Indels 24; Gaps 2;
QY 18 GTNSSPSTQNTSREVSVSSQLSEESTFYLCPPPGVSTVIRLEPPRKCPEPRKATWGE 77
DB 72 GTNASVEAGHATLRENLRDIALDGDATFFVCPPTGATVVQFPQPCPRAPHQONYTE 131
QY 78 GIALFKENISPYKFKVLYKNIQITTTTGTYYTQITNRYTDRTPSVSEIEITDLIDGK 137
DB 132 GIAVIFKENIAPYKFKATMYKDVTSQVWFGHRSQFMGIFEDRAPVPEEVDKINAR 191
QY 138 GRCSKARYLRNNVYVEAFDRDAGEKOVLLKPKSFNPESRAWHITNETYVVGSPWYIR 197
DB 192 GVCSTAKYVRNNMESTAFHRDDDESMDKAPAKAATRTSGWHTTDLKYNPSRIEAFHR 251
QY 198 TGTSVNCIVERMDARSVPFYSFAMANGDIANISPFYGLSPPEAAAPMGYPQDNFKOLD 257
DB 252 YGTTVNCIVEBEARSVPYDEFVLATGDFVYMSPFYGYR-DGAHAHTAAYADRPROVD 310
QY 258 SYFSDMLDKRKASLSPVKRNEPLTISHTFVGDWNAKPKTRVCSTKWKVEVTEMLRATVNGR 317
DB 311 GYERDLSTGRRASTPATRNLLATPKFTVGDWNAKPRPSVCTLTQWQEVDEMLRAEYGPS 370
QY 318 YRFMARELSATFISNTTEFPDNRITLIGOCIKREAAAEIOIFRTKYNDSHVKGHGVYFL 377
DB 371 FRFSSALSTFTTNRTEYALSVDLGDVCGREAREAVDRIFLRRYNGTCHVKVGOVQYLL 430
QY 378 ALGGFIVAYQVPLSKSLAHMYLRELMRDNRTDEMLDLVNNKHAIYKKNATSLSLRDRIR 437
DB 431 ATAGFLIAYQLLSNGLVELYVRELLRQGRP-----GDA 467
QY 438 NAPNRKITLDDTTAISTSSVQFAMLOFLYDHIQTHINDMFESRTATWCELQNLRELVLW 497
DB 468 ATPKSPADPPDVERIKTTSSVEFARLQFTYDHIQTHINDMFESRTATWCELQNLRELVLW 527
QY 498 EGKINPSATASATLGRVRAAKMGDVAASVSSCTAIDAESVTLQNSMRVITSTNTCYSRP 557
DB 528 EARLNPNAIASATVGRVRSARMGLDVMVAVSTCVPTPDNVMQNSMRVPRPGTCYSRP 587
QY 558 LVLFSGYNGQNIQOLGENNELLPTELEAVPCSAHRRYFLGSGYALFENYFVKWDAADIIQ 617
DB 588 LVSFYREEGGPLVEGOLGEDNEIRLERDALEPCTVGHRRYFTFGAGYVYFYEDYAYSHQLG 647
QY 618 AADIIQIASTFVELNLTLLEDEILPLSVYKKEELRDVGVDYAEVARRNQLHELK 672
DB 648 RADVTVTSTFINLNLTMLEDHEFVPLEVITRQEIKDSGLDLYTEVQRNQLHALR 702

RESULT 9
US-07-805-524-2
; Sequence 2, Application US/07805524
; GENERAL INFORMATION:
; APPLICANT: Babiuk, Lorne
; APPLICANT: Van Der Hurk, Sylvia
; APPLICANT: Zamb, Tim
; APPLICANT: Fitzpatrick, David
; TITLE OF INVENTION: Bovine Herpesvirus Type 1 Polypeptides
; TITLE OF INVENTION: and Vaccines
; NUMBER OF SEQUENCES: 10

```
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morrison & Foerster
;; STREET: 755 Page Mill Road
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/805,524
;; FILING DATE: 19911211
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gracey, Nancy J.
;; REGISTRATION NUMBER: 28,216
;; REFERENCE/DOCKET NUMBER: 293102000520
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-813-5600
;; TELEFAX: 415-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 933 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-07-805-524-2

Query Match 48.4%; Score 1698.5; DB 3; Length 933;
Best Local Similarity 50.0%; Pred. No. 3.3e-164;
Matches 331; Conservative 101; Mismatches 213; Indels 17; Gaps 5;

Qy 18 GTNSSPQNTVTSREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPPKCEPRKATWGE 77
Db 97 GDDAASPDNSTDVAARLAQAAGENSFRFCVCPSPGATVVRAPARCPCEPYGLGRNTE 156
Qy 78 GIALKFKENISPYKFKVLYKNIQTTWTGTYRQITNRYTDRTVPVSEETDLIDGK 137
Db 157 GIGVYKENIAPYTFKAYIYKKNIVITTWAGSYAAITNQYTDTRVPVNGEITDLVDK 216
Qy 138 GRCSKARYLRNNVYVEAFORDAGEKQVLLKPSKFNTPESRAWHTTNETYTVGSPMIY 197
Db 217 WRCLSKAEYLRSGKVVAFDRDDDPWEAPLKPALRSAPGVRGWHTTDDVYALGSAGLYR 276
Qy 198 TGTSVNCIVEMDARSVPYPYSYFAMANGDIANISPFYGLSPPEAAAEPMGYPDQNFOLD 257
Db 277 TGTSVNCIVEEARSVPYDPSFALSTGDIYMSPFYGLR-EGAHREHTSYSPERFOQIE 335
Qy 258 SYFSMDLDKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSMTKWKEVTEMLRATVNGR 317
Db 336 GYYRDMATGRLEKEPSRNFRLTQHTVAVDWPVKKNVCSLAKWEADEMLRDESRGN 395
Qy 318 YRFMARELSATFISNTTEFPDNRILGCQIKREAAAEIQIFRTKYNDSHVKVGHVQYFL 377
Db 396 FRFTARSLSATFVSDSHFALQNPVLSDCVTEEAEEAAVERVYRYNGTHVLSGSELYL 455
Qy 378 ALGFTVAYQPVLSKSLAHMYLRLEMRDNRDTEMLDVLNNKHAIYKKNATSLSLRDRIR 437
Db 456 ARGGFVAVFRPLMSNELAKYLQELARSNGTLEGLFAA----AAAPKFCP----RRARRAP 508
Qy 438 NAPNRKITL-----DDTTAIKSTSSVQFAMQLFDYDHIQTHINDMFSRIATAWCELQN 490
Db 509 SAPGGPGAANGPAGDGAGRGVTVSSAEFAALQFTYDHDQDHYNTMFSRLATSWCLLQN 568
Qy 491 RELYLWHEGKINPNSATGLRRVAAKMLGDVAAVSSCTAIDAESVTLQNSMRVITST 550
Db 569 KERALWAEAAKLNPSAAASALDRRAARMLGDAMAVTYCHELGEGRVFIENSAR--APG 626
Qy 551 NTCYSRLVLVFSYGENOGNIQGGQGENNELLPTLEAVEPCSANHRRYFLFGSGYALFENY 610

Db 627 GVCYSRPPVSFAFGNESEPEVGGQEDNELLPGRELYEPCETANKRYFRFGADVYVYENY 686
Qy 611 NFKWMDAADIAQIASTFEVLENLTLLEDEIREILPLSVYTKKEELRDVGVLDYAAVARNOLHE 670
Db 687 AYVRVPVLAELVISTFVDNLNLTVLEDEREFLPLEVYTRAEALADTGLDYSIQRRNQLHE 746
Qy 671 LK 672
Db 747 LR 748

RESULT 10
PCT-US03-11231-18
; Sequence 18, Application PC/TUS0311231
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.53801PC
; CURRENT APPLICATION NUMBER: PCT/US03/11231
; CURRENT FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
PCT-US03-11231-18

Query Match 48.2%; Score 1693; DB 1; Length 904;
Best Local Similarity 49.2%; Pred. No. 1.1e-163;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

Qy 20 NSSPSTQNTSVREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPPKCEPRKATWGBGI 79
Db 82 NATVAAGHATLRAHLREIKVENADAQFYVCPPPTGATVVOEQPRRCPTREPGQNYTGEI 141
Qy 80 AILFKENISPYKFKVLYKNIQTTWTGTYRQITNRYTDRTVPVSEETDLIDGKR 139
Db 142 AVFKENIAPYTFKATMIYKDVTVSQVNGHRYSQFMGIFEDRAPVPVEEVIDKINTKGV 201
Qy 140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAWHTTNETYTVGSPMIYRTG 199
Db 202 CRSTAKYVRNNMETTAFHRDDHETDMLKPAKVATRTSRGHTTDLKYNPSRVEAFHYG 261
Qy 200 TSVNCIVEEMDARSVPYPYSYFAMANGDIANISPFYGLSPPEAAAEPMGYPDQNFQDLSY 259
Db 262 TVNCIVEEVDARSVPYDFEVLATGDFVYMSPFYGR-EGSHTHTSYAADRQKQVDGF 320
Qy 260 FSDMLDKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSMTKWKEVTEMLRATVNGRYR 319
Db 321 YARDLTAKRATSPTRNLLTTPKFTVAVDWPVKRPVAVCTMTKQWEDMELRAEYGGSGFR 380
Qy 320 FMARELSATFISNTTEFPDNRILGCQIKREAAAEIQIFRTKYNDSHVKVGHVQYFLAL 379
Db 381 FSSDAISTTNTLNEYSLSRVDLGDICGRDAREIDRMFARKYNATHIKVGOQYQYLLAT 440
Qy 380 GGFIVAYQPVLSKSLAHMYLRLEMRDNRDTEMLDVLNNKHAIYKKNATSLSLRDRIRNA 439
Db 441 GGFLIAYOPLLSNLTALELYVRYMREQ-----DRK----PRNATP-----APLREA 482
Qy 440 PNRAITLDDTTAIAKSTSSVQFAMQLFDYDHIQTHINDMFSRIATAWCELQRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHQIRVNDMLGRIVAVWCELQHELTWNEA 539
Qy 500 IKINPSATASATGLRRVAAKMLGDVAAVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLAPNATASATVGRRRVARMGLGDVMAVSTCPVPADPNVIVQNSMRVSSRPGTCTYSRPLV 599
```

QY 560 LFSYGENOQN-IQOQIGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVA 618
Db 600 SFY-EDQGLIEGQGENNELRLTRDALEPCTVGHRRYFIFGGYVFEYAYSHQLSR 658
QY 619 ADIQIASTFVNLTLLEDEILPLSVYTKELRDVGVLDAEYAVARNOLHELK 672
Db 659 ADVTVTSTFDLNTLMEDEHFEVPLEVYTRHEIKDGLDYTEVQRNQLHDLR 712

RESULT 12
US-10-200-562-18
; Sequence 18, Application US/10200562
; GENERAL INFORMATION:
; APPLICANT: McGowen, Patrick
; APPLICANT: Hosken, Nancy A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C2
; CURRENT APPLICATION NUMBER: US/10/200,562
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-200-562-18

Query Match 48.2%; Score 1693; DB 28; Length 904;
Best Local Similarity 49.2%; Pred. No. 1.1e-163;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

QY 20 NSSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRLEPPKRCPEPRKATEWGEI 79
Db 82 NATVAAGHATLAHLREIKVENADAQFVCPPTGATVQFQPRCPTPEGQNYTEGI 141
QY 80 AILFENISPKFKVLYKNIQTTWTGTYRQITNRYTDRTVPSIEITDLIDGKR 139
Db 142 AVFKENIAPKFKATMYKDVTSQVWFGHRYSGFMGIFEDRAPVFEVIDKINTKV 201
QY 140 CSSKARYLRNVYVEAFDRDAGEQVLLKPSKFNTPESRAWHTTNETYTVWGSPIYRTG 199
Db 202 CRSTAKYVRNMETAFHRDDHETDMELKPAKVATRTSRGHTTDLKYNPSRVEAPHRYG 261
QY 200 TSNVCIVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAAEPMGYPODNFKQLDSY 259
Db 262 TTVNCIVEEDARSVPYDFEVLATGDFVYMSPFYGR-EGSHTHTSYAADRFQVDGF 320
QY 260 FSDMLDKRRKASLPVKNFLITSHFTVGMWMAKTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db 321 YARDLTTRKATSPTRNLLTTPKFTVAMDVWPKRPVACTMTKQEVDEMLRAEYGSFR 380
QY 320 FMARELSATISNTTEFPDNRILIGQCIRKREAAAEIQIFRTKYNDSHVKVGHVQYFLAL 379
Db 381 FSSDAISTFTTNLTETSLSRVDLGDGICGRDAREADRMFARKYNATHIKVGPQYLYAT 440
QY 380 GGFIVAYQVPLSKLAHMYRLMRDNRDTEMLDLVNKKHAIYKKNATSLSLRRDIRNA 439
Db 441 GGFLIAYQPLLSNTLAELYVREYREQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNKRTITDDTTAISTSSVQFAMQLFYDHIOTINDMFSRIATAWCELQNLRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNIHQHVNDMLGRJAVANCELQNLHELTLWNEA 539
QY 500 IKINPSATASATLGRVAAKMGDVAASVCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMGLDVMAVSTCVVPADPNVQNSMRVSRPGTCYSRPLV 599
QY 560 LFSYGENOQN-IQOQIGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVA 618
Db 600 SFY-EDQGLIEGQGENNELRLTRDALEPCTVGHRRYFIFGGYVFEYAYSHQLSR 658
QY 619 ADIQIASTFVNLTLLEDEILPLSVYTKELRDVGVLDAEYAVARNOLHELK 672
Db 659 ADVTVTSTFDLNTLMEDEHFEVPLEVYTRHEIKDGLDYTEVQRNQLHDLR 712

QY 560 LFSYGENOQN-IQOQIGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVA 618
Db 600 SFY-EDQGLIEGQGENNELRLTRDALEPCTVGHRRYFIFGGYVFEYAYSHQLSR 658
QY 619 ADIQIASTFVNLTLLEDEILPLSVYTKELRDVGVLDAEYAVARNOLHELK 672
Db 659 ADVTVTSTFDLNTLMEDEHFEVPLEVYTRHEIKDGLDYTEVQRNQLHDLR 712

RESULT 11
US-10-121-988-18
; Sequence 18, Application US/10121988
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowen, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

Query Match 48.2%; Score 1693; DB 27; Length 904;
Best Local Similarity 49.2%; Pred. No. 1.1e-163;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

QY 20 NSSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRLEPPKRCPEPRKATEWGEI 79
Db 82 NATVAAGHATLAHLREIKVENADAQFVCPPTGATVQFQPRCPTPEGQNYTEGI 141
QY 80 AILFENISPKFKVLYKNIQTTWTGTYRQITNRYTDRTVPSIEITDLIDGKR 139
Db 142 AVFKENIAPKFKATMYKDVTSQVWFGHRYSGFMGIFEDRAPVFEVIDKINTKV 201
QY 140 CSSKARYLRNVYVEAFDRDAGEQVLLKPSKFNTPESRAWHTTNETYTVWGSPIYRTG 199
Db 202 CRSTAKYVRNMETAFHRDDHETDMELKPAKVATRTSRGHTTDLKYNPSRVEAPHRYG 261
QY 200 TSNVCIVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAAEPMGYPODNFKQLDSY 259
Db 262 TTVNCIVEEDARSVPYDFEVLATGDFVYMSPFYGR-EGSHTHTSYAADRFQVDGF 320
QY 260 FSDMLDKRRKASLPVKNFLITSHFTVGMWMAKTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db 321 YARDLTTRKATSPTRNLLTTPKFTVAMDVWPKRPVACTMTKQEVDEMLRAEYGSFR 380
QY 320 FMARELSATISNTTEFPDNRILIGQCIRKREAAAEIQIFRTKYNDSHVKVGHVQYFLAL 379
Db 381 FSSDAISTFTTNLTETSLSRVDLGDGICGRDAREADRMFARKYNATHIKVGPQYLYAT 440
QY 380 GGFIVAYQVPLSKLAHMYRLMRDNRDTEMLDLVNKKHAIYKKNATSLSLRRDIRNA 439
Db 441 GGFLIAYQPLLSNTLAELYVREYREQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNKRTITDDTTAISTSSVQFAMQLFYDHIOTINDMFSRIATAWCELQNLRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNIHQHVNDMLGRJAVANCELQNLHELTLWNEA 539
QY 500 IKINPSATASATLGRVAAKMGDVAASVCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMGLDVMAVSTCVVPADPNVQNSMRVSRPGTCYSRPLV 599


```

Db 142 RACPDYKLGKNTGEGIAVIFENIAPYKFKANIYKNIIMTVMSSGYAVTNNRYTDRV 201
QY 124 PVSIEEITDLDGKGRCSKARYLRNNVYVEADRDAGEKQVLLKPSKFNTPESRAWHIT 183
Db 202 PVKVOEITDLDGKGRCSKARYLRNNVYVEADRDAGEKQVLLKPSKFNTPESRAWHIT 260
QY 184 NETYVWGSPIWYRTGYSVNCIVEEMDARSVPYPSYFAMANGDIANISPPYGLSPPEAAA 243
Db 261 NETYVWGSPIWYRTGYSVNCIVEEMDARSVPYPSYFAMANGDIANISPPYGLSPPEAAA 319
QY 244 EPMGYPODNFKQDYSFMDLDRKRRKASLPVKNRNLITSHFTVGVWDMAPKTRVCSMTKW 303
Db 320 EHTSYSSDRFQIEGYPIDLDYDTGA-PVSRNFELETPTHTVAVANNWTPKSGRVCTLAKW 378
QY 304 KEVTEMLRATVNGRYRFMAELSAFISNTTEFPDNRILGOCIKREAAEAEIQIFRTKY 363
Db 379 REIDEMPLMNT-GSYRFTAKTISATFISNTSQFEINRIRLGDCAATKAAEAIDRIYKSKY 437
QY 364 NDShVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRNDEMDDLNNKHAIYK 423
Db 438 SKTHIQTGTLETYLARGGLIAFRPMISNELAKLYINELARSNT--VVDL----SALLN 491
QY 424 KNATSLRLDRIRNAPNRKI-----TLDDTAIKSTSSVOFAMQLQFLYDHIOTH 473
Db 492 PSGETVQTRRSVPNSQHRHSRSTIEGGIETVNNASLLKTTSSVEFAMQLQFAYDIQAH 551
QY 474 INDMESRIATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDVAASVSCCTAI 533
Db 552 VNEMLSRATAWCTLQNRHVLWETLKNPGGVSMALERRVSARLLGDVAVATQCVNI 611
QY 534 DAESVTILQNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELLPTLEAVEPCSAN 593
Db 612 SSGHVYIQNSMRVTGSSTTCYSRPLVSFRALNDSEYIEGOLGENNELLVERKLIPECTVN 671
QY 594 HRRYFLFGSGYALFENYFVKMVDAAIOIASTFVELNLTLLEDREILPLSVYTKELRD 653
Db 672 NKRYFKFGADYVYFEDYAYKRVPLSEIELISAYV-IKSTLLEDREFLH-SSYTRAELED 729
QY 654 VGVLDYAEVARRNQLHELK 672
Db 730 TGPFDYSEIQRRNQLHALK 748

```

RESULT 15

```

US-09-521-738-14
; Sequence 14, Application US/09521738
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9
; CURRENT APPLICATION NUMBER: US/09/521,738
; CURRENT FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 943
; TYPE: PRN
; ORGANISM: Feline herpesvirus 1
US-09-521-738-14

```

```

Query Match 48.08; Score 1693.5; DB 19; Length 943;
Best Local Similarity 45.5%; Pred. No. 1.2e-162;
Matches 336; Conservative 129; Mismatches 193; Indels 81; Gaps 13;

QY 3 YERNICF-----FLIVLY-----GTNSSP 23
Db 22 YFRCQCFPILGATGSRHNGSSGLTRLARVSVFIWILVLPGRPVPEGSGSTSEQ 81
QY 24 STQNVTSREV-----VSSVOLSEER-----STFLCPPPVGVSTVIRLEPP 63

```

```

Db 82 PRRTVATPEVGVHQNQLQIPPICRYEALRASQIEANGPSTFYNCPPSGSTVVRLEPP 141
QY 64 RCPPEPRKATEWEGIAILFKENISPIYKFKVYKNIITOTTTGTGTQITNRYTDR 123
Db 142 RACPDYKLGKNTGEGIAVIFENIAPYKFKANIYKNIIMTVMSSGYAVTNNRYTDRV 201
QY 124 PVSIEEITDLDGKGRCSKARYLRNNVYVEADRDAGEKQVLLKPSKFNTPESRAWHIT 183
Db 202 PVKVOEITDLDGKGRCSKARYLRNNVYVEADRDAGEKQVLLKPSKFNTPESRAWHIT 260
QY 184 NETYVWGSPIWYRTGYSVNCIVEEMDARSVPYPSYFAMANGDIANISPPYGLSPPEAAA 243
Db 261 NETYVWGSPIWYRTGYSVNCIVEEMDARSVPYPSYFAMANGDIANISPPYGLSPPEAAA 319
QY 244 EPMGYPODNFKQDYSFMDLDRKRRKASLPVKNRNLITSHFTVGVWDMAPKTRVCSMTKW 303
Db 320 EHTSYSSDRFQIEGYPIDLDYDTGA-PVSRNFELETPTHTVAVANNWTPKSGRVCTLAKW 378
QY 304 KEVTEMLRATVNGRYRFMAELSAFISNTTEFPDNRILGOCIKREAAEAEIQIFRTKY 363
Db 379 REIDEMPLMNT-GSYRFTAKTISATFISNTSQFEINRIRLGDCAATKAAEAIDRIYKSKY 437
QY 364 NDShVKVGHVOYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRNDEMDDLNNKHAIYK 423
Db 438 SKTHIQTGTLETYLARGGLIAFRPMISNELAKLYINELARSNT--VVDL----SALLN 491
QY 424 KNATSLRLDRIRNAPNRKI-----TLDDTAIKSTSSVOFAMQLQFLYDHIOTH 473
Db 492 PSGETVQTRRSVPNSQHRHSRSTIEGGIETVNNASLLKTTSSVEFAMQLQFAYDIQAH 551
QY 474 INDMESRIATAWCELONRELVLWHEGKINPSATASATLGRVAAKMLGDVAASVSCCTAI 533
Db 552 VNEMLSRATAWCTLQNRHVLWETLKNPGGVSMALERRVSARLLGDVAVATQCVNI 611
QY 534 DAESVTILQNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELLPTLEAVEPCSAN 593
Db 612 SSGHVYIQNSMRVTGSSTTCYSRPLVSFRALNDSEYIEGOLGENNELLVERKLIPECTVN 671
QY 594 HRRYFLFGSGYALFENYFVKMVDAAIOIASTFVELNLTLLEDREILPLSVYTKELRD 653
Db 672 NKRYFKFGADYVYFEDYAYKRVPLSEIELISAYV-IKSTLLEDREFLH-SSYTRAELED 729
QY 654 VGVLDYAEVARRNQLHELK 672
Db 730 TGPFDYSEIQRRNQLHALK 748

```

Search completed: October 8, 2003, 17:25:01
Job time : 272.154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:07:16 ; Search time 38.6614 Seconds
(without alignments)
2800.682 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score:

Sequence: 1 MHYFRRNCIFFLIVLYGTN.....DVGVLDAEVARRNQLHELK 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

100% Processing. Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/FCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	3510	100.0	1086	9	US-09	US-09-147-052-4	Sequence 4, Appli
2	1693	48.2	904	9	US-09	US-09-894-998-18	Sequence 18, Appli
3	1693	48.2	904	12	US-10	US-10-200-552-18	Sequence 18, Appli
4	1693	48.2	904	12	US-10	US-10-237-551-18	Sequence 18, Appli
5	1693	48.2	904	15	US-10	US-10-121-988-18	Sequence 18, Appli
6	761	21.7	907	15	US-10	US-10-223-538-2	Sequence 2, Appli
7	745	21.2	830	14	US-10	US-10-055-364-45	Sequence 45, Appli
8	743	21.2	865	14	US-10	US-10-055-364-24	Sequence 24, Appli
9	737.5	21.0	793	14	US-10	US-10-055-364-55	Sequence 55, Appli
10	733	20.9	844	14	US-10	US-10-055-364-40	Sequence 40, Appli
11	721	20.5	808	14	US-10	US-10-055-364-38	Sequence 38, Appli
12	715.5	20.4	831	14	US-10	US-10-055-364-37	Sequence 37, Appli
13	699.5	19.9	823	14	US-10	US-10-055-364-39	Sequence 39, Appli
14	696	19.8	818	14	US-10	US-10-055-364-44	Sequence 44, Appli
15	694.5	19.8	829	14	US-10	US-10-055-364-46	Sequence 46, Appli

16	691	19.7	791	14	US-10-055-364-41	Sequence 41, Appl
17	683.5	19.5	792	14	US-10-055-364-42	Sequence 42, Appl
18	676	19.3	824	14	US-10-055-364-43	Sequence 43, Appl
19	669	19.1	359	14	US-10-055-364-48	Sequence 48, Appl
20	595.5	17.0	206	14	US-10-127-733-2	Sequence 2, Appl1
21	343.5	9.8	195	14	US-10-055-364-2	Sequence 2, Appl1
22	314	8.9	62	15	US-10-131-591A-5	Sequence 5, Appl1
23	314	8.9	456	9	US-09-147-052-2	Sequence 2, Appl1
24	314	8.9	456	12	US-09-901-572A-3	Sequence 3, Appl1
25	302	8.6	62	15	US-10-131-591A-6	Sequence 6, Appl1
26	113.5	3.2	433	12	US-10-288-930-83	Sequence 83, Appl
27	105.5	3.0	858	9	US-09-815-242-10894	Sequence 10894, A
28	103.5	2.9	1443	15	US-10-245-803-18	Sequence 18, Appl
29	101.5	2.9	662	15	US-10-128-714-8056	Sequence 8056, Ap
30	101.5	2.9	1164	10	US-09-870-123-1	Sequence 1, Appl1
31	101	2.9	1189	11	US-09-972-175-2	Sequence 2, Appl1
32	101	2.9	1189	11	US-09-972-175-4	Sequence 4, Appl1
33	101	2.9	1189	11	US-09-972-175-6	Sequence 6, Appl1
34	101	2.9	1189	11	US-09-972-175-8	Sequence 8, Appl1
35	101	2.9	1189	11	US-09-972-175-10	Sequence 10, Appl
36	101	2.9	1189	11	US-09-972-175-12	Sequence 12, Appl
37	101	2.9	1189	11	US-09-972-175-15	Sequence 15, Appl
38	101	2.9	1189	11	US-09-972-175-61	Sequence 61, Appl
39	101	2.9	1193	9	US-09-873-873-30	Sequence 30, Appl
40	101	2.9	1193	11	US-09-997-914-30	Sequence 30, Appl
41	101	2.9	1193	12	US-10-365-645-30	Sequence 86, Appl
42	101	2.9	2026	10	US-09-801-368-86	Sequence 11, Appl
43	100.5	2.9	716	14	US-10-036-492-11	Sequence 3, Appl1
44	100.5	2.9	1150	10	US-09-870-123-3	Sequence 28, Appl
45	99.5	2.8	1645	11	US-09-977-418-28	

ALIGNMENTS

RESULT 1

```

US-09-147-052--4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, NO. US20010014335A1oru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

```

Query Match	100.0%;	Score 3510;	DB 9;	Length 1086;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 672;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MHYFRNCIFFLIVILYGTNSSPSTQNTVTSREVYVSSVOLSEESTFYLCPPPVGVSTVIRL	60	
Db	1	MHYFRNCIFFLIVILYGTNSSPSTQNTVTSREVYVSSVOLSEESTFYLCPPPVGVSTVIRL	60	
QY	61	EPPRKCPERKATWEGEGIAILKENISPKYKFTVLYKNIQTTTWTGTYTYQIITNRYT	120	
Db	61	EPPRKCPERKATWEGEGIAILKENISPKYKFTVLYKNIQTTTWTGTYTYQIITNRYT	120	
QY	121	DRTPVSEIEITDLIDGKGRGSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSKFTPE\$RAW	180	

```

Db 121 DRTPIVSEIEITDLIDCKGRCSSKARYLNNVVEAFDRDAGEKQVLLKPSKENTPESRAW 180
QY 181 HTTNETYTVGSPWYRTGTGTVNCIVEEMDARSVPYFAMANGDIANISFFYGLSPPE 240
Db 181 HTTNETYTVGSPWYRTGTGTVNCIVEEMDARSVPYFAMANGDIANISFFYGLSPPE 240
QY 241 AAAPMGYPQDNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300
Db 241 AAAPMGYPQDNFKQDLSYFSDMDLDRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300
QY 301 TWKKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKREAAAEQIFR 360
Db 301 TWKKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKREAAAEQIFR 360
QY 361 TYNDSHVKGHVQVFLAGGIVAYQPVLSLAHMYLRELNRDNDMDLNNKHA 420
Db 361 TYNDSHVKGHVQVFLAGGIVAYQPVLSLAHMYLRELNRDNDMDLNNKHA 420
QY 421 IYKKNATLSRLRRDNRNAPNRKITDDTTAKSTSSVQFAMQLQFLYDHIOTHIHNDMFSR 480
Db 421 IYKKNATLSRLRRDNRNAPNRKITDDTTAKSTSSVQFAMQLQFLYDHIOTHIHNDMFSR 480
QY 481 IATAWCELONRELVLWHGEGIKNPSATASATLGRVAAKMGDVAVSSCTAIDAESVTL 540
Db 481 IATAWCELONRELVLWHGEGIKNPSATASATLGRVAAKMGDVAVSSCTAIDAESVTL 540
QY 541 QNSMRVITNTCYSRPLVFSYGENOGNITOGQGENNELPTLEAVEPCSAHRRYFLF 600
Db 541 QNSMRVITNTCYSRPLVFSYGENOGNITOGQGENNELPTLEAVEPCSAHRRYFLF 600
QY 601 GSGYALFENYFKVMVDAADIQIASTFVELNLTLEDREILPLSVYTKBELRDVGVDYA 660
Db 601 GSGYALFENYFKVMVDAADIQIASTFVELNLTLEDREILPLSVYTKBELRDVGVDYA 660
QY 661 EVARRNQLHEK 672
Db 661 EVARRNQLHEK 672

```

```

RESULT 2
US-09-894-998-18
; Sequence 18, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-09-894-998-18

```

```

Query Match 48.2%; Score 1693; DB 9; Length 904;
Best Local Similarity 49.2%; Pred. No. 2.1e-157;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

QY 20 NSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPVPGVSTVIRLEPPRCPCPEKATGEI 79
DB 82 NATVAAGHATLAHLREIKVENADAFYVCPPTGATVQVFEQPRCPTPEQONTYEG 141
QY 80 ALLFKNISPKFKVLYKNIQTWTGTYRTNRYTDRTPVSEIEITDLIDCKGR 139
DB 82 NATVAAGHATLAHLREIKVENADAFYVCPPTGATVQVFEQPRCPTPEQONTYEG 141
QY 80 ALLFKNISPKFKVLYKNIQTWTGTYRTNRYTDRTPVSEIEITDLIDCKGR 139

```

```

Db 142 AVFKEKNIAPYKFKATMYKDVTVSQVWFGRHYSQFMGIFEDRAPVPFEEVIDKINTKV 201
QY 140 CSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAWHTTNETYTVGSPWYRTG 199
Db 202 CRSTAKYVRNNMETAFHRDDHETDMELKPAKVATFTRSGWHTTDLKYNPVRSEAFHYG 261
QY 200 TSVNCIVEEMDARSVPYFAMANGDIANISFFYGLSPPEAAEPMGYPQDNFKQDLSY 259
Db 262 TVNCIVEEMDARSVPYFAMANGDIANISFFYGLSPPEAAEPMGYPQDNFKQDLSY 320
QY 260 FSDMDLDRKASLPVKRNFLITSHFTVGDWAPKTRVCSMTWKVEYTEMLRATVNGRYR 319
Db 321 YARDLTTKARATSTPTTNLLTTPFTVANDWVPAPVCTWKQOEDEMLRAEYGGSPR 380
QY 320 FMARELSATFISNTTFFDPNRIILGQCIRKREAAAEQIFRKYNDSHVKGHVQVFLAL 379
Db 381 FSSDAISFTTNTCYSRPLVFSYGENOGNITOGQGENNELPTLEAVEPCSAHRRYFLF 440
QY 380 GGFIVAPQVLSLAHMYLRELNRDNDMDLNNKHAIIYKKNATLSRLRRDNRNA 439
Db 441 GGFLAYQPLSLNTLAELYVREYREQ-----DRK----PRNATP-----APLEA 482
QY 440 PNKKTLDTTAKSTSSVQFAMQLQFLYDHIOTHIHNDMFSRIATAWCELONRELVLWHG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYHNRVNDMLGRIVAAWCELONRELVLWHG 539
QY 500 IKNPSATASATLGRVAAKMGDVAVSSCTAIDAESVTLQNSMRVITNTCYSRPLV 559
Db 540 RKLNPATASATVGRVSRMGLDVMAYSTCVVPADNVIVONSRRVSRGTCYSRPLV 599
QY 560 LFSYGENOGNITOGQGENNELPTLEAVEPCSAHRRYFLGSGYALFENYFKVMVDA 618
Db 600 SPRY-EDOGPLEGOLGENNELPTLEAVEPCSAHRRYFLGSGYALFENYFKVMVDA 658
QY 619 ADIQAISTFVELNLTLEDREILPLSVYTKBELRDVGVDYAFAARRNQLHEK 672
Db 659 ADVTTVSTFIDLNITMLEDHEFVPLEVYTRHEIKDGLDYTEVQRNQLHDLR 712

```

```

RESULT 3
US-10-200-562-18
; Sequence 18, Application US/10200562
; Publication No. US20030165819A1
; GENERAL INFORMATION:
; APPLICANT: McGowan, Patrick
; APPLICANT: Hosken, Nancy A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C2
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-200-562-18

```

```

Query Match 48.2%; Score 1693; DB 12; Length 904;
Best Local Similarity 49.2%; Pred. No. 2.1e-157;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

QY 20 NSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPVPGVSTVIRLEPPRCPCPEKATGEI 79
DB 82 NATVAAGHATLAHLREIKVENADAFYVCPPTGATVQVFEQPRCPTPEQONTYEG 141
QY 80 ALLFKNISPKFKVLYKNIQTWTGTYRTNRYTDRTPVSEIEITDLIDCKGR 139
Db 142 AVFKEKNIAPYKFKATMYKDVTVSQVWFGRHYSQFMGIFEDRAPVPFEEVIDKINTKV 201
QY 140 CSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAWHTTNETYTVGSPWYRTG 199

```

```

Db      202 CRSTAKYVNNMHTAFHRDDHETDMLKPAKVATRTSRGWHTTDLKYNPSRVEAFHRYG 261
QY      200 TSVNCIVEEMDARSVPYSYFAMANGDIANTISPFYGLSPPEAAAPMGYPQDNFKOLDSY 259
Db      262 ITVNCIVEVDARSVPYDEFLATGDFVYNSPFYGYR-EGSHTHTSTAAADRFRKQVDGF 320
QY      260 FMDLDDKRRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db      321 YARDLTTKARATSPTRNLLTTPKFTVADWVWPKRPACTMTKWQEVDEMLRAEYGGSGFR 380
QY      260 FMDLDDKRRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db      321 YARDLTTKARATSPTRNLLTTPKFTVADWVWPKRPACTMTKWQEVDEMLRAEYGGSGFR 380
QY      320 FMARELSATFISNTTFFDPNRIILGQCICKREAAAEQIFRTKYNDSHVKVGHVQYFLAL 379
Db      381 FSSDAISTFTTNLTYSLSRVDLGDICIGRDAREIDRMFARKYNATHIKVQPOYYLAT 440
QY      380 GGFIVAYOPVLSKSAHMYLRELMRDNTDMDLVLNNKHAIYKKNATSLSLRLRDINA 439
Db      441 GGFLIAYOPVLSKSAHMYLRELMRDNTDMDLVLNNKHAIYKKNATSLSLRLRDINA 439
QY      440 PNKRTITLDDTTAIKTSSVOFAMLOFLYDHIOTHIHNDMFRIATACELQNLRELVLWHEG 499
Db      483 PSANASVE---RIKTTSSIEFARLOFTYNNHQRHNDMLGRITAVANCELOHLETLWNEA 539
QY      500 IKINPSATASATLGRRAAKMLGDVAASVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db      540 RKLNPATASATVGRVSARMGLDVMVSTCVVPADNVIVQNSMRVSRPCTCYSRPLV 599
QY      560 LFSYGENQGN-IOGOLGENNELLPLEAVEPCSAHRRYFLFGSGYALFENTNFVKMVD 618
Db      600 SPRY-EDQGPLEGOLGENNELRLTRDALEPCTVGHRRYFFGGYVYFEEYAYSHQLSR 658
QY      619 ADIQIATSFVLENLTLLEDERELPLSVYTKBELRDVGVLDAEVARRNOLHELK 672
Db      659 ADVTTVSTFIDNLTMLEDHEFVPLEVYTRHEIKDGLDYTEVQRRNQLHDLR 712

```

RESULT 4

```

US-10-237-551-18
; Sequence 18, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Joseph M.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-237-551-18

```

```

Query Match      48.2%; Score 1693; DB 12; Length 904;
Best Local Similarity 49.2%; Pred. No. 2.1e-157;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

```

```

QY      20 NSSPSTQNTVSREVSVSSVOLSEESTFYLCPPPVGSTVIRLEPPKRCPEPRKATWEGEI 79
Db      82 NATVAAGHATLAHLREIKVENADAQFYVCPPTGATVVQFQPRCPTRPBGQNTYEGI 141
QY      80 ATLFKENTISPKFKVLYKKNIIQTWTGTYRQITNRYTDRTVPSIEEITDLIDGKR 139
Db      142 AVVFENIAPYKFKATMYKDVTSQVWFGHYSQFMGIFEDRAVPVPEEVDKINTKGV 201
QY      140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAWHTTNETYTVWGSPIYRTG 199
Db      202 CRSTAKYVNNMHTAFHRDDHETDMLKPAKVATRTSRGWHTTDLKYNPSRVEAFHRYG 261
QY      200 TSVNCIVEEMDARSVPYSYFAMANGDIANTISPFYGLSPPEAAAPMGYPQDNFKOLDSY 259

```

```

Db      262 ITVNCIVEVDARSVPYDEFLATGDFVYNSPFYGYR-EGSHTHTSTAAADRFRKQVDGF 320
QY      260 FMDLDDKRRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db      321 YARDLTTKARATSPTRNLLTTPKFTVADWVWPKRPACTMTKWQEVDEMLRAEYGGSGFR 380
QY      320 FMARELSATFISNTTFFDPNRIILGQCICKREAAAEQIFRTKYNDSHVKVGHVQYFLAL 379
Db      381 FSSDAISTFTTNLTYSLSRVDLGDICIGRDAREIDRMFARKYNATHIKVQPOYYLAT 440
QY      380 GGFIVAYOPVLSKSAHMYLRELMRDNTDMDLVLNNKHAIYKKNATSLSLRLRDINA 439
Db      441 GGFLIAYOPVLSKSAHMYLRELMRDNTDMDLVLNNKHAIYKKNATSLSLRLRDINA 439
QY      440 PNKRTITLDDTTAIKTSSVOFAMLOFLYDHIOTHIHNDMFRIATACELQNLRELVLWHEG 499
Db      483 PSANASVE---RIKTTSSIEFARLOFTYNNHQRHNDMLGRITAVANCELOHLETLWNEA 539
QY      500 IKINPSATASATLGRRAAKMLGDVAASVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db      540 RKLNPATASATVGRVSARMGLDVMVSTCVVPADNVIVQNSMRVSRPCTCYSRPLV 599
QY      560 LFSYGENQGN-IOGOLGENNELLPLEAVEPCSAHRRYFLFGSGYALFENTNFVKMVD 618
Db      600 SPRY-EDQGPLEGOLGENNELRLTRDALEPCTVGHRRYFFGGYVYFEEYAYSHQLSR 658
QY      619 ADIQIATSFVLENLTLLEDERELPLSVYTKBELRDVGVLDAEVARRNOLHELK 672
Db      659 ADVTTVSTFIDNLTMLEDHEFVPLEVYTRHEIKDGLDYTEVQRRNQLHDLR 712

```

RESULT 5

```

US-10-121-988-18
; Sequence 18, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

```

```

Query Match      48.2%; Score 1693; DB 15; Length 904;
Best Local Similarity 49.2%; Pred. No. 2.1e-157;
Matches 322; Conservative 114; Mismatches 194; Indels 24; Gaps 7;

```

```

QY      20 NSSPSTQNTVSREVSVSSVOLSEESTFYLCPPPVGSTVIRLEPPKRCPEPRKATWEGEI 79
Db      82 NATVAAGHATLAHLREIKVENADAQFYVCPPTGATVVQFQPRCPTRPBGQNTYEGI 141
QY      80 ATLFKENTISPKFKVLYKKNIIQTWTGTYRQITNRYTDRTVPSIEEITDLIDGKR 139
Db      142 AVVFENIAPYKFKATMYKDVTSQVWFGHYSQFMGIFEDRAVPVPEEVDKINTKGV 201
QY      140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAWHTTNETYTVWGSPIYRTG 199
Db      202 CRSTAKYVNNMHTAFHRDDHETDMLKPAKVATRTSRGWHTTDLKYNPSRVEAFHRYG 261

```

QY 200 TSVCNIVEEMDARSVPYGFYFAMANGDIANISPFYGLSPPEAAABPMGYQPDNFKQLDSY 259
Db 262 TVNCIVEEDARSVPYDFEVLATGDFYMGSPFYGYR-EGSHTESYAADREFQVDGF 320
QY 260 FMSDLKRRKASLPVKRNELIISHFTVGVNDWAPKTRVCSMTKKEVTEMLRATVNGRYR 319
Db 321 YARDLTKARATSPTRNLITTPKFTVANDWPKRPVAVCTMTKQWQVDEMLRAEYGGSPR 380
QY 320 FMARELSAFISNTTFEDPNRIILGOCIKREAEAAEQIFRTKINDSHVKVHVQVFLAL 379
Db 381 FSSDAISTFTTTLNLTYSLSURVDGDCIGRDAREADRMFAKYNATHKVGQPOPYLAT 440
QY 380 GGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMIDLNNKHAIYKKNATLSLSLRDIRNA 439
Db 441 GGFIIAYQPLLSNTLAELYVREYMRQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNKRIITLDDTTAKTSSTSSVOFAMQLQFLYDHIQTHINDMFSRIATANCELNRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHQIRHVNDMLGRIAVANCELNHELTLWNEA 539
QY 500 IKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMLGDVMAVSTCVPAVDNVIQNSMRVSRPGCYSRPLV 599
QY 560 LFSYGENQGN-IOGOLGENNELPTLEAVBPCCSANHRRYFLFGGYALPENYFVKMVA 618
Db 600 SFYR-EDQGPLIEQOLGENNELRLTRDALEPCTVGHRRYEFIFGGYVFEYAYSHQLSR 658
QY 619 ADIOIATSTFVELNLTLEDEILPLSVYTKRELDRGVGLDYAEVARRNOLHELK 672
Db 659 ADVTTSTFIDLNTIMLEDHEFVLEPTEVITRHEINDSGLLDYTEVQRNQLHDLR 712

RESULT 6
US-10-223-538-2
; Sequence 2, Application US/10223538
; Publication No. US20030120060A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Karl, Csaba
; TITLE OF INVENTION: No. US20030120060A1el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/223,538
; FILING DATE: 19-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-223-538-2

Query Match 21.7%; Score 761; DB 15; Length 907;
Best Local Similarity 29.3%; Pred. No. 1.8e-65;
Matches 212; Conservative 122; Mismatches 275; Indels 114; Gaps 18;

QY 8 CTFEFLIVILYGTNSPST-----QNTVSRREVSS----- 36
Db 11 CVNLICIVCLGAAVSSSTRGTSATHSHSSHTSAHSRSGSVSRVTSSTQSVSHGVNET 70
QY 37 -----VOLSEESTFVLCPPPGVSTVIRLEPPKCPKPEPKATE-WGEGTAILFK 84
Db 71 IYNTTLKYGDVVGVNTKYPYRVCSMAOQTDLIRERNIVCTSMKPINEDLDEGLVYVK 130
QY 85 ENISPYKFKVTLIYKNII--QTTWTGTYTQITNRYDRTPTVSEETDLIDGRC-S 141
Db 131 RNVAHTFKRVYQVLTFRSYAVIHTYLLGSN--TEYVAPPMEIHH-INSHSQYS 187
QY 142 SKARYLRNNVYEAEDRDAGE-KQVLLKPSKENTPESRAWHTTNETYTVMGSPWLYRTGT 200
Db 188 SYSRVTAGTVFV-AYHRDSYENKTMQMLPDDTSNTHSTRIVTKOOWHSRGSTWLYRET 246
QY 201 SYNCIVEEMDARSVPYGFYFAMANGDIANISPFYGLSPPEAAABPMGYQPDNFKQLDSY 260
Db 247 NLNCMTIITARSKYPYHEFATSGDVVDISPFYNGCTRNA-----SYF 290
QY 261 SMDLDK-----RRKASLPVVR--NFLTSHFTVGVNDWAPKTRVCSMTKK 304
Db 291 GENADKFFIFPNYTIIVSDFGRPNSALETHRLVAFLEADSVISWDQDEKNVTCOLTFWE 350
QY 305 EVTEMLRATVNGRYREFMARELSATFISNTEFPDNRILGQCICKREAAAEIQIFRTKN 364
Db 351 ASERTIRSEADSYHFSSAKWTATFLSKQOEVNMSDAL-DCVRDEAINKLQOIFWTSYN 409
QY 365 DSHVKVGHVOYFALGGFIVAYOPVLSKSLAHMYLRELMR-DNRTDEMIDLNN--KHAI 421
Db 410 QTYEKYGNVSFVETGGLVVFQGIQKSLV-----ELERLANRSS--LNLTHNRTKST 462
QY 422 YKKNATSLSLRLDRDIRNAPNRKITLDDTTAIKTSSVOFAMQLQFLYDHIQTHINDMFSRI 481
Db 463 DGNATHLSNM-----ESVHNLVYTAQLQFTTDLRGYNRALAQI 502
QY 482 ATAWCELQNRRELVLWHEGKINPFSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTIQ 541
Db 503 AEAWCYDQRTLEVFVKELSKINPISAILSAIYNKPIAAREFMGDLGLASCVTINQTSVKVL 562
QY 542 NSMRVITSTNTCYSRPLVLSYGENQGNTOGQGENNELPTLEAVBPCCSANHRRYFLFG 601
Db 563 RDMNVKESPGRCYSRPVFIENFANSSVYQYCGLEDNEILLGNHRTTECOLPSLKTFIAG 622
QY 602 SGVALFENYFVKMVAADIQIASTFVELNLTLEDEILPLSVYTKRELDRGVGLDYAE 661
Db 623 NSAYEYVDYLFKRMIDLUSSISTVDSMTALDIDPLENTDFRVLEYLSQKELRSSNVFDLEE 682
QY 662 VAR 664
Db 683 IMR 685

RESULT 7
US-10-055-364-45
; Sequence 45, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

```
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Alcelaphine herpesvirus
US-10-055-364-45

Query Match      21.2%; Score 745; DB 14; Length 830;
Best Local Similarity 29.4%; Pred. No. 5.9e-64;
Matches 197; Conservative 110; Mismatches 310; Indels 54; Gaps 14;

QY 8 CIIFFLIYUG-----TNSPSTQNTVTSREVSVSSVQLSEEST-----FYL 48
Db 9 CAFLIFAVLKNVFCQPTSSSEVEDVPEANTVSDNIIRQORNTAKGIHSDPSAFPERV 68
QY 49 CPPPVGSTVIRLEPPKCPKPRKATEWEGGIALFKENISPKFKVTLTKYKNIQTTWT 108
Db 69 CSASNTGDIIFRQTSKCFN-TKDKHEHNEGILLFKENIVPVYFKYRKIVTTSTIYN 127
QY 109 GTTYRITNRYTDRTPVSTEEITDLIDGKRCSSKARYLNRYVVEAFDRDAGEKOVLLK 168
Db 128 GIYADAVTQHFVSKSPVIE-TRMDTIYQCNLSLDVTVGGNLLVYTDNDGSNWTVDLQ 186
QY 169 PSKFNTPESRAHHTTNETYVWGSPWI---YRGTSVNCIVEEMDARSVPFYSYFAMWG 225
Db 187 PVDGLNSVRRVHSQPEIHAEPG--WLLGGYRRRTTVNCEVETDARAVPPFRYFTINIG 244
QY 226 DIANISPFGLSPPEAAAPMGYPQDNFKQDSYFSMDLDRKASLPVKRNFELIISHT 285
Db 245 DTIEMSPFWSKAWNE--TEFSGEPDRTLVAKDYRVVDYKFRGTQPGQHTRIPIVDKEET 302
QY 286 VGDWDAKTRVCSMTKWEVTEMLRATVNGRYRFMARELSATFISNTTEFDNRIILGQ 345
Db 303 LSWAQOFRNISYCRWAHWSFDNAIKTEHGSLSLHFVANDITASFTPTNQ---TREVLK 359
QY 346 --CIKREAAEABIEQIFRTYKNSHVGHVQYFALGGFIVAYQPVLKSLAHMYLRELM 403
Db 360 HVCNNTIESELKSRL-AKVNDTHSPNGTAQYLYLTNGLLLVWQPLVQOKLLD----- 411
QY 404 RDNRTDMDLVNKKHAIYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVQFAML 463
Db 412 ----AKGLLDVAVK-----QONTTNTTTRSRQRSSVSGIDDDVYTAEST--ILLTQI 460
QY 464 QFYLDHIQTHINDMFRIATAWCALQNRRLVWHEGIKINPSATASATLGRVAAKMLGD 523
Db 461 QFAYDPLRAQINNVLSELSAWCQREHRAHSLMNLKINPTSVMSIYGRPVSAKRIGD 520
QY 524 VAAVSCSTAIDASVTLQNSMRV--ITSNTCYSRPLVLFSGYNGENGTQOGLGENNELL 581
Db 521 VISVCHVVVDQDSVSLHRSMPVRGDKTHECYSRPPVTFKFINDSHLYKGLQGVNNEIL 580
QY 582 PTLAEVPCSANHRRYFLGSGVALFENYFKWVDAADIQIASTFVELNLTLLEDRIL 641
Db 581 LTTTAVIEICHENTHEHYFOGNNMNYFYKNYRHKVTMPGVDAVTLDTFMVLNLTLVENIDFQ 640
QY 642 PLSVYTKELR 652
Db 641 VIELYSREKR 651
```

RESULT 8

US-10-055-364-24

; Sequence 24, Application US/10055364

```
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence of porcine gamma herpesvirus gpB g
US-10-055-364-24

Query Match      21.2%; Score 743; DB 14; Length 865;
Best Local Similarity 28.8%; Pred. No. 1e-63;
Matches 193; Conservative 121; Mismatches 295; Indels 60; Gaps 13;

QY 17 YGTNSSP-----STQNTVTSREVSVSSVQLSEESTFVLCPPVPYRCSASGVGVDFRQTDHVCPPASD 71
Db 51 YGTHDSSHGCGNENRDSSEQNKNIYSPSTPYRCSASGVGVDFRQTDHVCPPASD 110
QY 72 ATEWGGIALFKENISPKFKVTLTKYKNIQTTWTGTYRITNRYTDRTPVSTEEIT 131
Db 111 MVH-SEGILLIYKQNIIPFMRVRYKRVVTVTSTVNGIYSDSITNQHTFYKSIEPWE-T 168
QY 132 DLIDGKRCSSKARYLNRYVVEAFDRDAGEKOVLLKPSKFNTPESRAHHTTNETYT--- 188
Db 169 ERMDTIYQCNLSRLTGGNLLVYDVRDINWTVFQPDGVTPDKRYGSGQELYLEPG 228
QY 189 -VWGSPWIYGTGTSVNCIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPEAAAPMG 247
Db 229 WFWGS---YRRRTTVNCELMDMFARSNPDPDFVTATGDTVEMSPW--SGEDHENKMH 283
QY 248 YPDQNFQDLSYFSMDLDRKASLPVKRNFELIISHTFVGDWDAKTRVCSMTKWEVT 307
Db 284 EKPWFVSVINNYKVVQYQNRGTVPGLKTRIFLDREYVTLSEKHLKNNMTCPLTLWKAFY 343
QY 308 EMLRATVNGRYRFMARELSATETS-----NTTEFDNRIILGOCIKREAAEABIEQIF 359
Db 344 NGIQTSHSGSYHFVANDITASFTTSKEDMKFEFTTY-----HCLNEEIKAEIEKKY 394
QY 360 RTKYNDSHVKVGHVQYFALGGFIVAYQPVLKSLAHMYLRELMRDNRTDMDLVNKKH 419
Db 395 -AKVNSTHSGYGLKFKYDKDGLLVWQPLIQ-----NRLDANKNLN 436
QY 420 AITKKNATSLRLRRDIRNAPNRKITLDDTTA---TKSTSSVQFAMQFYLDHIQTHIND 476
Db 437 -----NETYSRRSRQAEESTDPMMEMTNGAGGEYSSENSITVAQYAYDNLIRINN 491
QY 477 MFSRIATAWCALQNRRLVWHEGIKINPSATASATLGRVAAKMLGDVAAVSCSTAIDAE 536
Db 492 ILEDLSKAWCQREHRAALVWNLKSNPTSVMSIYGRPVSAKRIGDIVSVCIVVDQT 551
QY 537 SVTLQNSMRVIT--STNTCYSRPLVLFSGYNGENGTQOGLGENNELLPTLEAVPCSANHR 595
Db 552 SVSLHSLRLLSADEKCFSPRPVTFKFNNDSTIYKGLGVNNEILLTLYLTQCENTE 611
QY 596 RYFLGSGVALFENYFKWVDAADIQIASTFVELNLTLLEDRILPLSVYTKELRDVG 655
Db 612 YFQAKTDMYIYKNYEHLKTVPLSSITLDTLTFALNFTLLENVDFKVIELYTRDKRLSN 671
QY 656 VLDYAEVAR 664
```

Db 672 VFDIETMER 680

RESULT 9

US-10-055-364-55

; Sequence 55, Application US/10055364

; Publication No. US20020155433A1

; GENERAL INFORMATION:

; APPLICANT: Patience, Clive

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

; FILE REFERENCE: 61750-379

; CURRENT APPLICATION NUMBER: US/10/055,364

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/612,204

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US/60/142,736

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: US/60/168,532

; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 55

; LENGTH: 793

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Portion of Acetaphine herpesvirus.

US-10-055-364-55

Query Match 21.0%; Score 737.5; DB 14; Length 793;

Best Local Similarity 30.8%; Pred. No. 3e-63;

Matches 189; Conservative 102; Mismatches 288; Indels 35; Gaps 12;

QY 46 FYLCPPPVGSGTVIRLEPPKCPPEPRKATEWEGIALEFENISPKFKVTLIYKNIQTT 105

Db 12 FVCSASNIGDIFRQTSKSCPN-TKDEHNEGILLIKENIVPVFVKRYKIVTST 70

QY 106 TWGTGTYRQITNRYDTRPTVSTEEITDLDGKGRCSKARYLRNNVYVEAFDRAGEKQV 165

Db 71 IYNGIYADAVTNGHVSFVSPIYE-TRRMDTIYQINSLDVTVGNGLLVYTDNDSNMTV 129

QY 166 LKPSKFNTPESRAWHHTNETVTVMGSPWI---YRTGTSVNCIIVEMDARSVPYSYFAM 222

Db 130 DLOPVGLSNVRRYHSQPEIHAEPG--WLLGGYRRRTVNCVEVETDARAVPPERYFIT 187

QY 223 ANGDIANISPFYGLSPPEAAEPMPGYPQDNFKOLDYSFMDLDRKASLPVKRNFILTS 282

Db 188 NIGDTIEMSPFWSKANNE--TEFSGEPDRTLTVAKYRVVDYKFRGTOPQGHTRIFVDKE 245

QY 283 HFTVGDWAPKTRVCSMTKKEVTEMLRATVNGRYRFEMARELSATFISNTTEFDPNRII 342

Db 246 EYTLWAQDFRNISYCRWAHWSFDNAIKTEHGKSLHFVANDITASFYTPNQ---TREV 302

QY 343 LQO--CIKREAAAIQIFRTKYNDSHVGVQVFLALGGFIVAYQPVLSKSLAHMYLR 400

Db 303 LGRHVCLNNTIESELASRL-AKVNDTHSPNGTAQYVLTNGLLVWQPLVQCKLLD--- 357

QY 401 ELMRDNRDMLDLVNKKHAIYKKNATSLRLRDRINAPNRKIITLDDTTAKTSVQVQF 460

Db 358 -----AKGLDAVKK-----QONTTITTTTTSRRQRSSVSSGIDDDVYTAEST--ILL 403

QY 461 AMLQFLYDHIOTINDMFBSRIATWCELQNLRELVLWHEGIIKINPSATASATIGRRVAAKM 520

Db 404 TQIQFAYDTLRAQINNVLELSRAWCEBQRASLMLNELSKINPTSVMSIYGRPVSAKR 463

QY 521 LGDVAAYSSCTAIDAEVTLQNSMRV--IISTWCYSRPLVLFPSYGENGNTOGQGENN 578

Db 464 IGDVISVSHCVVQDQSVLSHRMVRPGDRDKTECYSRPVPTEKFIINDSHLYKQGLGVNN 523

QY 579 ELLPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVAADIQIASTFVELNLTLLEDR 638

Db 524 EILLTTTAVTECHENTBHYFQGGNNMYFYKNYRHVKTMPPGVGDVATLDTFMVLNLTIVENI 583

QY 639 EILPLSVYTKKEELR 652

Db 584 DFQVIELYSREEKR 597

RESULT 10

US-10-055-364-40

; Sequence 40, Application US/10055364

; Publication No. US20020155433A1

; GENERAL INFORMATION:

; APPLICANT: Patience, Clive

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

; FILE REFERENCE: 61750-379

; CURRENT APPLICATION NUMBER: US/10/055,364

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/612,204

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US/60/142,736

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: US/60/168,532

; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 40

; LENGTH: 844

; TYPE: PRT

; ORGANISM: Bovine herpesvirus 4

US-10-055-364-40

Query Match 20.9%; Score 733; DB 14; Length 844;

Best Local Similarity 29.4%; Pred. No. 9.3e-63;

Matches 197; Conservative 109; Mismatches 305; Indels 60; Gaps 16;

QY 19 TNSPSTQNV-TSREVVSVVOLSEESTFY---LCPPPVGS-TVIRLEPPKCPPEPRKAT 73

Db 62 STKSPSTDNGTSTPTTPTVTDTSKFNFKYKVCSSSSGSELFDPDLDDQTCPD-TKDK 120

QY 74 EWGGIALLKFNKISPKFKVTLIYKNIQTTWTGTYRQITNRYDTRPTVSEIEITDL 133

Db 121 KHVEGILLVLKKNIVPIFKVRYKRTATSVYRGWSQAQVNRDDISRAIPYNEIS-M 179

QY 134 IDGKGRCSKARYLRNNVYVEAFDRAGEKOVLLKPSKFNTPESRAWHHTNETIYTVWG-S 192

Db 180 IDRTYHCFSAMATVINGILNTYIDRSENKSVLPQVPVAGLTENINRYFSQPLIYAEFGWF 239

QY 193 PWIYRTGTSVNCIIVEMDARSVPYSYFAMANGDIANISPF-----YGLSPPEAAAE 244

Db 240 PGTYRVRTTVNCEVDDMTARSVEPTHTTALGDITIELISFCHNNSOCITGNSTSRDATK 299

QY 245 PMGYPDQNFKOLDSYFSMDLDRKASLPVKRNFILTSHTFTVGDWAPKTRVCSMTKWK 304

Db 300 V--WIEENHQTVDY-----ERRGHPTKDKRIELKDEEVTISWKAEDRERAIQDFVIWK 350

QY 305 EVTEMLRATVNGRYRFEMARELSATFISNT---TEFDNRIILGOCIKREAAAEQIFRT 361

Db 351 TFPRAIQTIHSHSFHFVANEVTSFELTSNOBELTGRNGTEIL-NCMNSTINETLEETVK- 408

QY 362 KYNDSHVGVQVFLALGGFIVAYQPVLSKSLAHMYLRRLMRDNRDMLDLVNNKHAI 421

Db 409 KFNKSHIRDGVEKYKTKNGGLFLIWAQMKPLNLSEHTNYTIERNKT-----GNK--- 458

QY 422 YKKNATSLRLRDRINAPNRKIITLDDTTAKTSVSVQFAMQLQFLYDHIQTHINDMFSEI 481

Db 459 -----SROKRSV-----DTKTFQAGKGLSTAQVQAYAYDHLRTSMHILEEL 499

QY 482 ATAWCELQNLRELVLWHEGIIKINPSATASATIGRRVAAKMGDVAASVCTAIDAESVTLQ 541

Db 500 TKTWCROKKNMLMYELSKINPVSVMIAIYKGPVAVKAMGDAPVSCINVDQASVNIH 559

QY 542 NSMRVITSTNTCYSRPLVLFPSYGENGNTOGQGENNELLPTEAVEPCSANHRRYFLFG 601

Db 560 KSMRT-DDPKVCYSRPLVTFKFPVNSTATFRGQLGRNEILLTNTHTVETCRPTADHYFPVK 618

QY	477	MFSRTATWCELQNLRELVLWHEGKIKNPSATASNTLGRVAAKMGDVAVSSCTAIDAE	536
DB	475	VLEELSRACREQVDNKLWMLKSIKPIPTSVMTAIGRPVSAKFGDAISVTEICINVDOS	534
QY	537	SVTLQNSMRVITSTNTCYSRPLVLFSTYGENQNGTQGLGENNELLPTEAVEPSCSAHRR	596
DB	535	SVNIKSLRT-NSKDDVCARPLVTFKELINSNLTFTGOLGARNELIITLNQVETCKDTCEH	593
QY	597	YFLGSGVALPENYNFVKWMDAADITQASTFVELNLTLEDREILPLSVVYTKPELR	652
DB	594	YFTTRFETVLVKDYALRTINTDITSLNTFIALNFIQALNFIQDKAILEYSSAEKR	649

```

RESULT 13
US-10-055-364-39
; Sequence 39, Application US/10055364
; Publication No. US2002015433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Murine herpesvirus 68
US-10-055-364-39

```

```

QY      512 LGRVAAKMLGDAVYAASCTAIDAESVTLQNSMRVITSTNCTYSRPLVLFSYGENOGNIQ 571
Db      508 YKPVAARYVGDAISVTDCLIVDOSSVNIHQSLQHDKTCTYSRPRVKTFKFINSTDPLT 567
QY      572 GOLGENNELIPTLEAVEPCSANRRRYFLFGSGYALFENYNFKVMVAADIQIASTFFVELN 631
Db      568 GOLGPRKEILSNWTNIETCKDESEHFVIGEYYIYYKNYIFEEKLNLSSTATLDTTFTALN 627
QY      632 LTLLREIREILPSVYTKEE 650
Db      628 ISFIENIDFKTVELYISTE 646

RESULT 14
US-10-055-364-44
; Sequence 44, Application US/10055364
; Publication No. US20020155433AI
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/169,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRNT
; ORGANISM: Equine herpesvirus 5
US-10-055-364-44
```


THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:40 ; Search time 18.888 Seconds
(without alignments)
1505.341 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672

Perfect score: 3510

Sequence: 1 MHYFRNCIFFLIVLYGTN.....DVGVLDAEVARNQLHELK 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3510	100.0	865	1	US-07-803-633A-13
2	1898.5	54.1	879	1	US-08-220-151-2
3	1898.5	54.1	879	1	US-08-220-151-3
4	1898.5	54.1	879	1	US-08-413-118-2
5	1898.5	54.1	879	1	US-08-413-118-3
6	1898.5	54.1	879	1	US-08-413-118-106
7	1898.5	54.1	879	3	US-08-473-446-2
8	1898.5	54.1	879	3	US-08-473-446-3
9	1898.5	54.1	879	3	US-08-473-446-106
10	1786.5	50.9	1041	1	US-08-220-151-4
11	1786.5	50.9	1041	1	US-08-413-118-4
12	1786.5	50.9	1041	3	US-08-473-446-4
13	1726	49.2	868	1	US-08-220-151-7
14	1726	49.2	868	1	US-08-413-118-7
15	1726	49.2	868	3	US-08-804-439A-21
16	1726	49.2	868	3	US-08-473-446-7
17	1726	49.2	868	3	US-08-720-229-21
18	1719.5	49.0	913	3	US-09-232-468A-2
19	1719.5	49.0	913	3	US-09-784-984B-49
20	1715.5	48.9	913	1	US-08-220-151-6
21	1715.5	48.9	913	1	US-08-413-118-6
22	1715.5	48.9	913	3	US-08-473-446-6
23	1699	48.4	891	1	US-08-042-747A-6
24	1697.5	48.4	913	6	5196516-8
25	1697.5	48.4	933	1	US-08-682-847-2
26	1697.5	48.4	980	1	US-08-220-151-5
27	1697.5	48.4	980	1	US-08-413-118-5

28	1697.5	48.4	980	3	US-08-473-446-5	Sequence 5, Appl
29	1693	48.2	904	4	US-09-894-998A-18	Sequence 18, Appl
30	1692	48.2	904	3	US-08-632-537-2	Sequence 2, Appl
31	1692	48.2	904	5	PC-T-US96-05316-2	Sequence 2, Appl
32	1692	48.2	904	6	5244792-3	Patent No. 5244792
33	1683.5	48.0	943	3	US-08-911-321-4	Sequence 4, Appl
34	1679	47.8	885	1	US-08-042-747A-8	Sequence 8, Appl
35	1674	47.7	885	3	US-08-804-439A-23	Sequence 23, Appl
36	1674	47.7	885	3	US-08-720-229-23	Sequence 23, Appl
37	1663	47.4	904	3	US-08-632-537-1	Sequence 1, Appl
38	1663	47.4	904	5	PC-T-US96-05316-1	Sequence 1, Appl
39	1660	47.3	903	3	US-08-804-439A-22	Sequence 22, Appl
40	1660	47.3	903	3	US-08-720-229-22	Sequence 22, Appl
41	1657	47.2	903	1	US-08-220-151-8	Sequence 8, Appl
42	1657	47.2	903	1	US-08-413-118-8	Sequence 8, Appl
43	1657	47.2	903	3	US-08-473-446-8	Sequence 8, Appl
44	1654	47.1	904	6	5244792-4	Patent No. 5244792
45	761	21.7	907	1	US-08-349-006-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-803-633A-13

Query Match 100.0%; Score 3510; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 3.9e-315;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVLYGTNNSPSTQNTSVREVSVSVQLSEESTFYLCPPVGVSTVRL 60

```

Db 1 MHYERRNCIFLIIVILGTHNSPSTQNTSREVSVSSVQLSEESTFYLCPPVSTVRL 60
Qy 61 EPPKRCPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTYRQITNRYT 120
Db 61 EPPKRCPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTYRQITNRYT 120
Qy 121 DRTDVSIEETDLDGGRGSSKARYLRNNVYVEAFDRDAGEKQVLKPKSFNTPESRAW 180
Db 121 DRTDVSIEETDLDGGRGSSKARYLRNNVYVEAFDRDAGEKQVLKPKSFNTPESRAW 180
Qy 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSIPFYGLSPPE 240
Db 181 HTTNETYTVMGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSIPFYGLSPPE 240
Qy 241 AAEPMGYPQDNFKQDLSYFMDLDRKKRSLPVKRNFLITSHFTVGMWNAKPTRVCSM 300
Db 241 AAEPMGYPQDNFKQDLSYFMDLDRKKRSLPVKRNFLITSHFTVGMWNAKPTRVCSM 300
Qy 301 TKKKEVTEMLRATVNGRYRFMARELSATFISNTTEFPDNRILIGOCIKREAAAEQIFR 360
Db 301 TKKKEVTEMLRATVNGRYRFMARELSATFISNTTEFPDNRILIGOCIKREAAAEQIFR 360
Qy 361 TKYNDSHVKVGHQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDVLNNKHA 420
Db 361 TKYNDSHVKVGHQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDVLNNKHA 420
Qy 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVQFAMQLFYDHIOTINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVQFAMQLFYDHIOTINDMFSR 480
Qy 481 IATAWCQLQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSCAIDAESVTL 540
Db 481 IATAWCQLQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSCAIDAESVTL 540
Qy 541 QNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPCSANHRRYFLF 600
Qy 601 GSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKELRDVGVLDDYA 660
Db 601 GSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKELRDVGVLDDYA 660
Qy 661 EVARRNOLHELK 672
Db 661 EVARRNOLHELK 672

```

```

RESULT 2
US-08-220-151-2
; Sequence 2, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-220-151-2

Query Match 54.1%; Score 1898.5; DB 1; Length 879;
Best Local Similarity 52.2%; Pred. No. 2.9e-166;
Matches 356; Conservative 120; Mismatches 187; Indels 19; Gaps 4;

Qy 9 IFFLIIVL-----YGTNSSPSTQNTSREVSVSSVQL-SEESTFYLCPP 51
Db 8 IFFLIIVL-----YGTNSSPSTQNTSREVSVSSVQL-SEESTFYLCPP 51
Qy 52 PVGSTVRLPPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTT 111
Db 52 PVGSTVRLPPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTT 111
Qy 68 PSGSTVRLPPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTT 127
Db 68 PSGSTVRLPPEPRKATEWGEGIAILFKENISPYKFKVTLVYKKNIIQTWTGTT 127
Qy 112 YRQITNRYTTPVSIIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLKPSK 171
Db 112 YRQITNRYTTPVSIIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLKPSK 171
Qy 128 YAVITNRYTDRVPIGVPEITELIDRRGMCLSKADYIRNNYEFTAFKDEDPREVHLKPSK 187
Db 128 YAVITNRYTDRVPIGVPEITELIDRRGMCLSKADYIRNNYEFTAFKDEDPREVHLKPSK 187
Qy 172 FNTPESAWHTTNETYTVMGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSI 231
Db 172 FNTPESAWHTTNETYTVMGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSI 231
Qy 188 FNTPGSGWHTVNDTYTKIGSGFYHSGTSVNCIVEEMDARSVPYSYFAMANGDIANSI 247
Db 188 FNTPGSGWHTVNDTYTKIGSGFYHSGTSVNCIVEEMDARSVPYSYFAMANGDIANSI 247
Qy 232 PFYGLSPPEAAEPMGYPQDNFKQDLSYFMDLDRKKRSLPVKRNFLITSHFTVGMWDA 291
Db 232 PFYGLSPPEAAEPMGYPQDNFKQDLSYFMDLDRKKRSLPVKRNFLITSHFTVGMWDA 291
Qy 248 PFGLR-DGAHTEVYSYSTDQFQIEGYYPIDLDRLQLGAPVSRNFLTQHVTYVANNV 306
Db 248 PFGLR-DGAHTEVYSYSTDQFQIEGYYPIDLDRLQLGAPVSRNFLTQHVTYVANNV 306
Qy 292 PKTRVCSMTKWKVEMLRATVNGRYRFMARELSATFISNTTEFPDNRILIGOCIKREA 351
Db 292 PKTRVCSMTKWKVEMLRATVNGRYRFMARELSATFISNTTEFPDNRILIGOCIKREA 351
Qy 307 PKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATFISDTTQFDIDRVKLSDCAKREA 366
Db 307 PKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATFISDTTQFDIDRVKLSDCAKREA 366
Qy 352 EAAIEQIFRTKYNDSHVKVGHQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDTEM 411
Db 352 EAAIEQIFRTKYNDSHVKVGHQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDTEM 411
Qy 367 IEALDKIYKKYKNTHTQTELEYLARGGFITAFRPMISNELAKLYNELVSRNRTVDL 426
Db 367 IEALDKIYKKYKNTHTQTELEYLARGGFITAFRPMISNELAKLYNELVSRNRTVDL 426
Qy 412 LDLVNNKHAIYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVQFAMQLFYDHIQ 471
Db 412 LDLVNNKHAIYKKNATSLRLRRDIRNAPNRKITLDDTTAKSTSSVQFAMQLFYDHIQ 471
Qy 427 KSLINPSVGGARKRSRVEENKRSKRIEGGIENNNSTIIKTSSVHFAMQLFYDHIQ 486
Db 427 KSLINPSVGGARKRSRVEENKRSKRIEGGIENNNSTIIKTSSVHFAMQLFYDHIQ 486
Qy 472 THINDMFSRATWCELQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSC 531
Db 472 THINDMFSRATWCELQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSC 531
Qy 487 SHYNEMLSRATWCELQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSC 546
Db 487 SHYNEMLSRATWCELQRELVLWHEGKINPSATASATLGRVAAKMGDVAASVSC 546
Qy 532 AIDAESVTQNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPC 590
Db 532 AIDAESVTQNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPC 590
Qy 547 NISGSSVFIQNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPC 606
Db 547 NISGSSVFIQNSMRVITSTNTCYSRPLVLFSGYNGQNIQGGQGENNELLPTLEAVEPC 606
Qy 591 SANHRRYFLGSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKEE 650
Db 591 SANHRRYFLGSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKEE 650
Qy 607 TANHRRYFLGSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKEE 666
Db 607 TANHRRYFLGSGYALFENTFYKMWDAADIQASTFVFNELNLLDREILPLSVYTKKEE 666
Qy 651 LRDVGVLDDYAEVARRNOLHELK 672
Db 651 LRDVGVLDDYAEVARRNOLHELK 672
Qy 667 LEDTGLDYSEIQRRNOLHELK 688
Db 667 LEDTGLDYSEIQRRNOLHELK 688

```

```

US-08-220-151-3
; Sequence 3, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOSIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtiss, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-3

```

Query Match	54.1%	Score	1898.5;	DB 1;	Length	879;			
Best Local Similarity	52.2%;	Pred. No.	2.9e-166;						
Matches	356;	Conservative	120;	Mismatches	187;	Indels	19;	Gaps	4;

QY	9	IFPLIVL-----YGTNSPSTQNTVSREVSSVOL-SEESTFVLCPP	51
DB	8	IFFIIITLIICDPTTPESTINPNLHNLSTPKPSTDIDREILRESQIESDDTSTFYMCPP	67
QY	52	PVGSTVIRLEPPKRCPEPRKATEGWGGIATLFKENISPKYFKVTLVYKNIQTWTGTT	111
DB	68	PGSGLVRLLEPPRACPNYKLGKFTGIAVIFENISPKFKANIYKNIITVWSGST	127
QY	112	YROITNRYDTRTPVSEIEITDLIDGRCRCSKARYLRNNVYVAFDRDAGEKQVLLKPSK	171
DB	128	YAVITNRYDTRVPIGVPEITELIDRRGMLCSKADYIRNNVEYFAFDKDEDPREVHLKPSK	187
QY	172	FNTPESSRAHHTTNETTVWGSPLHIYTGTSVNCIVEMDARSVPFYSYFAMANGDTANIS	231
DB	188	FNTPGSRGWHVTNDTYTKIGGSGFYHSGTSVNCIVEDARSVPYDPSFAISTGDIHMS	247
QY	232	PFYGLSPPEAAAPMGYPQDNFKQDSYFSMDLDRKKAASLPVKRNFLLTSHETVGDWA	291
DB	248	PFTGLR-DGAHTYEIYSITDRFOOIGYIPDIDTLQLGAPVSRNFLTQHVTYVANNV	306
QY	292	PKTRVCSMTKKKEVTEMLRATVNGYRFRMARELSATFISNTTFEPNRIILQCICKREA	351
DB	307	PKIREVCTLAKWEIDELIRDEYKGSYRTAKSISATFFISDTQFIDRDKVLSDCAKREA	366
QY	352	EAAEQIFRTKYNDSHVKGVHVOYFIALGGFTIVAYQFVLSKSLAHMYELRMDRNDTDEM	411

```

367 IEADIKYKKYKXTHQTQGELETYLARGGFIIFRPMISNELAKLYINELVRSNRTVDL 42
412 LDLVNKHAIYKKKATSLRRURDIRNAPNRKITLDDTTATKSTSSVOFAMQLQFLYDHIQ 471
427 KSLNPSVRGGARKRSVEENKRSKRNTIEGGIENVNNSTIINKTSSVHFAMQLQFAYDHIQ 486
472 THINDMFRIATACEQLONRELVLWHEGIKINPSATASATIGRRVAAKMLGDVAAVSSCT 531
487 SHVNEMLSRATACNLQNKERTLWNEVKMLNPTSVASVAMDQVRSARMIGDVLAVTQCV 546
532 AIDAESVTLQNSMRVISTNTCYSRPLVLFSGYENQGN- IQQLGENNELLPLEAVEPC 590
547 NISGSSVFQNSMRVLGSTTTCYSRPLISFKALENSTNYIEQQLGENNELLYVERKLIIEC 606
591 SANHRRYFLFGSGYALFENYFVKMVDADIQIASTFVELNLTLEDREILPLSVYTKKE 650
607 TANHKRYFKGADYVFENYAYVRKVPFLNEIEMISAYVDLNTLTLEDREFLEFVYTRAE 666
651 LRDVGVLIDYAEVARNQLHELK 672
667 LEDTGLLDYSEIQRRNQLHALK 688

RESULT 4
US-08-413-118-2
; Sequence 2, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-413-118-2

```

Query Match 54.1%; Score 1898.5; DB 1; Length 879;
Best Local Similarity 52.2%; Pred. No. 2.9e-166;
Matches 356; Conservative 120; Mismatches 187; Indels 19; Gaps 4;

QY 9 IFFLIVIL-----YGTNSSPSTQNTVSREVVSQVL-SEESTFYLCPP 51
Dd 8 IFFIYIILICDPTTPESTINPLNHNHLSLTPKPTSDIREILRESQIESDDTSFYMCPP 67
QY 52 PVGSTVIRLEPRKCPKPRKATEWEGIAILFKENISPYKFKVLYKNIITTTWTGTT 111
Dd 68 PSGLTLVRLEPRACPNYKLGKNTGEGIAVFKENISPYKFKANIKYNIITTTWGSST 127
QY 112 YRQITNRYTDRTPVSIIEITLDIGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSK 171
Dd 128 YAVITNRYTDRVPGLVPEITELIDRGMCLSKADYIRNNYEFTAFDDEDPREVHLKPSK 187
QY 172 FNTPESSRAWHNTETNYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANIS 231
Dd 188 FNTPGSRGHWNTDYTKIGSGFYHSGTSVNCIVEEDARSVPYSYFAMANGDIANIS 247
QY 232 PFYGLSPPEAAEPMPGPDQNFKOLDSYFSDMLDKRKASLPVKRNFELITSHFTVGDWA 291
Dd 248 PFFGLR-DGAHTEYISYSTRFOQIEGYPIIDLTLRLQGLGAPVSRNFLTTHQVTVANWV 306
QY 292 PKTTRVCSMTKWKVTEMLRATVNGRYRFHARELSATFISNTTFEDPNRIILGOCIKREA 351
Dd 307 PKIREVCTLAKWREIDELIRDEYKGSYRFTAKSISATFISDTTQFDIDRVKLSDCAKREA 366
QY 352 EAAIEQIFRTKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEM 411
Dd 367 IEADIKYKKYKNTHTIQTGELETYLAGGFIIAFRPMISNELAKLYNELVRSNRITVDL 426
QY 412 LDLYNNKHAIYKKNATSLSLRLRRDIRNAPNRKITLDDTTAITSQVQFAMQLFYDHIQ 471
Dd 427 KSLNPSVRGARKRRSRVEENKRSKRNIIEGGENNSTIITKTSSVHFAMQLFYDHIQ 486
QY 472 THINDMFSRTATANCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAASVST 531
Dd 487 SHVNEMLSRITATACNLQNKERTLWNEVMKLNPTSVASVAMDQVRSARMLGDVLAQTQCV 546
QY 532 AIDAESVTLQNSMRVITSTNCTYSRPLVLSYGENOQN-IOGQLGENNELLPTELEAVEPC 590
Dd 547 NISGSSVFIQNSMRVLGTTCTYSRPLISFKALENSTNYIEGQLGENNELLIVERKLEIPC 606
QY 591 SANHRRYFLGSGYALFENFVKWDAADTQIATSTFVNLTLLEDEILPLSVYTKKEE 650
Dd 607 TANKRYFKGADYVYENAYVRKVPNLNETEMISAYVDNLITLLEDEFLPLEVYTRAE 666
QY 651 LRQVGLDYAEVARRNQLHELK 672
Dd 667 LEDTGLLDYSEIORRNQLHALK 688

RESULT 5
US-08-413-118-3
Sequence 3, Application US/08413118
Patent No. 5688920
GENERAL INFORMATION:
APPLICANT: PAOLETTI, ENZO
APPLICANT: LIMBACH, KEITH J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,118

FILING DATE: 29-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 879 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-413-118-3
Query Match 54.1%; Score 1898.5; DB 1; Length 879;
Best Local Similarity 52.2%; Pred. No. 2.9e-166;
Matches 356; Conservative 120; Mismatches 187; Indels 19; Gaps 4;
QY 9 IFFLIVIL-----YGTNSSPSTQNTVSREVVSQVL-SEESTFYLCPP 51
Dd 8 IFFIYIILICDPTTPESTINPLNHNHLSLTPKPTSDIREILRESQIESDDTSFYMCPP 67
QY 52 PVGSTVIRLEPRKCPKPRKATEWEGIAILFKENISPYKFKVLYKNIITTTWTGTT 111
Dd 68 PSGLTLVRLEPRACPNYKLGKNTGEGIAVFKENISPYKFKANIKYNIITTTWGSST 127
QY 112 YRQITNRYTDRTPVSIIEITLDIGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSK 171
Dd 128 YAVITNRYTDRVPGLVPEITELIDRGMCLSKADYIRNNYEFTAFDDEDPREVHLKPSK 187
QY 172 FNTPESSRAWHNTETNYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANIS 231
Dd 188 FNTPGSRGHWNTDYTKIGSGFYHSGTSVNCIVEEDARSVPYSYFAMANGDIANIS 247
QY 232 PFYGLSPPEAAEPMPGPDQNFKOLDSYFSDMLDKRKASLPVKRNFELITSHFTVGDWA 291
Dd 248 PFFGLR-DGAHTEYISYSTRFOQIEGYPIIDLTLRLQGLGAPVSRNFLTTHQVTVANWV 306
QY 292 PKTTRVCSMTKWKVTEMLRATVNGRYRFHARELSATFISNTTFEDPNRIILGOCIKREA 351
Dd 307 PKIREVCTLAKWREIDELIRDEYKGSYRFTAKSISATFISDTTQFDIDRVKLSDCAKREA 366
QY 352 EAAIEQIFRTKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDNRNRTDEM 411
Dd 367 IEADIKYKKYKNTHTIQTGELETYLAGGFIIAFRPMISNELAKLYNELVRSNRITVDL 426
QY 412 LDLYNNKHAIYKKNATSLSLRLRRDIRNAPNRKITLDDTTAITSQVQFAMQLFYDHIQ 471
Dd 427 KSLNPSVRGARKRRSRVEENKRSKRNIIEGGENNSTIITKTSSVHFAMQLFYDHIQ 486
QY 472 THINDMFSRTATANCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAASVST 531
Dd 487 SHVNEMLSRITATACNLQNKERTLWNEVMKLNPTSVASVAMDQVRSARMLGDVLAQTQCV 546
QY 532 AIDAESVTLQNSMRVITSTNCTYSRPLVLSYGENOQN-IOGQLGENNELLPTELEAVEPC 590
Dd 547 NISGSSVFIQNSMRVLGTTCTYSRPLISFKALENSTNYIEGQLGENNELLIVERKLEIPC 606
QY 591 SANHRRYFLGSGYALFENFVKWDAADTQIATSTFVNLTLLEDEILPLSVYTKKEE 650
Dd 607 TANKRYFKGADYVYENAYVRKVPNLNETEMISAYVDNLITLLEDEFLPLEVYTRAE 666
QY 651 LRQVGLDYAEVARRNQLHELK 672
Dd 667 LEDTGLLDYSEIORRNQLHALK 688

```

307 PKIREVCTAKWREIDRIIRDEYKGSYRFTAKSISATFISDTQFDDRVKLSDCAKREA 366
352 EAIEQIFRKYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDDEM 411
367 IEAIDKIKYKKYKNTKHTQGTGLETYLARGGFIIAFRPMISNELAKLYINELVRSNRTVDL 426
412 LDLVNKKHAIYKKKNTSLSRLLDIRNAPNRKITLDDPTAKTSSSVQFAMQLQFLYDHQ 471
427 KSLNPSVRGAGARKRRSVENKRSKRNIIEGGIENVNNTIIKTSSVHFAMQLQFAYDHQ 486
472 THINDMFRIATAMCELQNLRELVLWHEGIKINPSATASATLGRRVAAKMLGDVAASVSC 531
487 SHVNEMLSRITATWCNQLNKERTLWNEVMKLNPTSVASVAMDQRVSARMIGDLVAVTQCV 546
532 AIDAESVTYONSRVITSTWTCYSRPLVLFSGYGENQGN-IOGQNGENNELPTLEAVEPC 590
547 NISGSSVFIOQNSMRVLSTTTCYSRPLTSFKALENSTNYIEGQNGENNELVERKLIIEP 606
591 SANHRRVFLFGSGVALEPNFVKWVDAADTQIASTFVELNUTLLEDEREILPLSVYTREE 650
607 TANHKRYFKGADYVFENTAYVRKVPNLNEIEMISAYVDLNTLLEDEREFLPLEVYTRAE 666
651 LRDVGVLDAEVARRNQLHELK 672
667 LEDTGLLDYSIEIQRNRQLHALK 688

RESULT 7
US-08-473-446-2
; Sequence 2, Application 05/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; ; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; ; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-473-446-2

```

Query Match	54.1%	Score 1898.5	DB 3	Length 879
Best Local Similarity	52.2%	Pred. No. 2.9e-166		
Matches 356	Conservative 120	Mismatches 187	Indels 19	Gaps 4
QY	9	IFFLIVIL-----YGTNSPSTQNTVTSREVSSVOL-SEESTSYFLC	PP	51
			:	
Db	8	IFFIYLIICDPTTPESTINPLMHHLSLPKPTSDDIRILRESQIESDDTSTYMC	PP	67
			:	
QY	52	PVGSVIRLEPPKPCPRKATEMGEGIALFKENISPYKFKVLYLYKKNIOTIWTGT	T	111
			:	
Db	68	PGSFTVLEPPRACPNYKLGKNTGEGIAVFNENISPYKFRANIYKNIITTVSGST	T	127
			:	
QY	112	YQIINRYTDRTPVSIIBEITDLIDKGRCSKARYLRNNVYVFAFDRAGEKOVLLKPSK		171
			:	
Db	128	YAVINRYTDRVPICGVEITELIDRGMCLSKADYIRNNYEFATFDKEDPREVHLKPSK		187
			:	
QY	172	FNTPESSRAWHHTNHYTVWGSPLYRPGTGVNCVIEEMDARSVPYSYFAMANGDIANS		231
			:	
Db	188	FNTPGSRGWHVTNDTYTKIGSGYHSGTGVNCVIEVDARSVPYDVSFAISTGDIHMS		247
			:	
QY	232	PYIGLSPPEAAEPNGYPODNFKOLDYSFMDLKRKRKASLPVKRNFLLTSHTFVGWDA		291
			:	
Db	248	PFEGLR-DGAHTEYISYDFRQIEGYYPIDLDLRLQLGAPVSRNFLTQHVTVANNV		306
			:	
QY	292	PXTTRVCMTKWKEVTEMLRATVNGRYRFMARELSATFISNTTFEPDNRILILGOCIKREA		351
			:	
Db	307	PKRIREVCTLAKWREIDEIIRDEYKSGYRFTAKSISATFISDTTQFIDIRVKLSDCAKREA		366
			:	
QY	352	EAAIEQIERTYNDSHKVGHVQVFLAGGFIVAYQPVLSKSLAHMYLRELARDNRDTEM		411
			:	
Db	367	IEADIKYIKKYNKTHQTGTLETYLAGGFIIFAPRMISNELAKIYINELVRSNRTVDL		426
			:	
QY	412	LQLVNKHAIYKKNATLSRLRDRDIRAPNRKKTLDTTTAKTSVSSQFAMLQFLYDHQI		471
			:	
Db	427	KSLNPSVYSGGARRKRSVEENKSKRIEGEINVNNSITIKTSSVHFAMLQFAYDHQI		486
			:	
QY	472	THINDWFRIATANCELRNRELVLWHBGIKINPISATASATLGRVRAAKMLGDVAASVST		531
			:	
Db	487	SHVNEMLSRITATWCNLQNKERTLWNEVMKLNTPSVASVAMDQVRSARMLGDVLAVTCV		546
			:	
QY	532	AIDAESVTLQNSMRVINTSNFCYSRPLVLSYGENQGN-IGQGLGNNELLFTLEAVRCP		590
			:	
Db	547	NTSGSSVYIQNSMRVLGSTTTCYSRPLISPKALENSTNYLEGQLGNNELLVERKLIEPC		606
			:	
QY	591	SANHRRYFLGSGVALFENYFVKWDAADIIQASTFVELNLTLLDREILPLSVYTKEE		650
			:	
Db	607	TANHRRYKFGADYVYFENYAYVKVPVPLEMEISAMSYVDNLNLTLLDREFLPLEVYTRAE		666
			:	
QY	651	LKDVGVLDAEYVARNQLHLEK		672
			:	
Db	667	LEDTGILDYSEIQRNOLHALK		688
			:	

```

RESULT 8
US-08-473-446-3
; Sequence 3, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```



```

Db      607  TANHRRYKFGADYYFENYAVYKRVPLNEIEMISAYVDLNTLLSDREFLPLEVYTRAE 666
QY      651  LRDVGLDYAEVARRNQLHELK 672
Db      667  LEDTGLDYSQIQRNQLHALK 688

RESULT 9
US-08-473-446-106
; Sequence 106, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
US-08-473-446-106

Query Match          54.1%; Score 1898.5; DB 3; Length 879;
Best Local Similarity 52.2%; Pred. No. 2.9e-166;
Matches 356; Conservative 120; Mismatches 187; Indels 19; Gaps 4

QY      9  IFPLVIL-----YGTNSPSTONVTSREWVSSVOL-SEESTFVLCPP 51
Db      8  IFFIYTLICDPTTPESTINPLNHNLSKPKTSDDIRILRESQIESDDTSTFFWCPP 67
QY      52  PVGSTVIRLEPPKCPERKATWEGEGTAILFKENISPYFKVTLYYKNIQTWTGTT 111
Db      68  PGSGTLVRLPEPPRACPNYKLGKNTGEGIAVFKENISPYAFKANIYYKNIITVWSGST 127
QY      112 YQIOTNRYTDRFTPSIEETDLIDGKRCSSKARYLRNNYVFAFDORAGEKOVLLKPSK 171
Db      128 YAVITNRYTDRVPICVPEITELIDRRGMCLSKADYIRNNYEFTAFOKDEDPREVHLKPSK 187
QY      172 FNTGPSRAWHHTNEHYTVWGSFWIYRGTCTSYNCIVEEMDARSVPFYSYFAMANGDIANIS 231
Db      188 FNTGPSRGWHTVNDYTKIGGSGFYHGTCTSYNCIVEEVDARSVPYFSDYFAISTGDIHMS 247

```

MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-220-151-4

Query Match 50.9%; Score 1786.5; DB 1; Length 1041;
Best Local Similarity 45.4%; Pred. No. 8.3e-156;
Matches 343; Conservative 126; Mismatches 185; Indels 101; Gaps 5;

QY 18 GTNSSPSTQNTVTSREVSSVOLSEEE-----STFYLCPPPVGTVIRLEPPKCPPEPRKAT 73
DB 93 GTPPKPTDPTDMSDREALRASQIEANGSTFTYMCPPSGTVVRLPPRACPDYKLGK 152

QY 74 EWGECIALFKENISPVYKFKVLYKNIIQTTTGTGTYRQITNRYTDRTPVSEIEITDL 133
DB 153 NFTGIAVIFKENIAPYKFKANIYKNIIMTVWSSGYATVNNRYTDRVPKVOEITDL 212

QY 134 IDGGRSSKARYLRNNVYVEAFDRDAGEKQVLKPSKFNTPESR----- 178
DB 213 IDRGMCILSKADYVRNNYQFTAFDRDEDPRELPLKPSKFNTPQSRGWHYKFKATVYVKD 272

QY 179 ----- 178

QY 273 VIVSTAWAGSSYQITNRYADRPVSEITDIDKFGKCSKATYVRNNHKEAFNEDEK 332

QY 179 -----AWHTTNETVTVMGSPWYRTGTSVNCIVEEMDARSVPYYSF 220

QY 333 NPQDMPLIASKYNSVSGKAWHTTNETYTKGAAGFHSGTSVNCIVEEDARSVPYDSF 392

QY 221 AMANGDANTISPFYGLSPPEAAAPMGYPQDNFKQSDYSFMDLDRKRSKASLPYKRNFLI 280

QY 393 AISTGDVIHMSPPFGLR-DGAHVEHTSYSSDRFQIQIEGYPIDLDTLRLQLGAPVSRNLE 451

QY 281 TSHTFTGWDWAPKTRVCSMTKWEVTEMLRATVNGRYRPMARELSATFISNTTFEPDNR 340

QY 452 TPHYVAVNWTPKCGRVCTIAKWEIDEMLRDEYQGSYRFTVTKTISATFISNTSQFEINR 511

QY 341 IILGQCICKREAAAEIQRFTKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLR 400

QY 512 IRLGDCATKEAAEAIDRIYKSKYKTHIQGTLETYLARGGFLIAPRPMISNELAKLYIN 571

QY 401 ELMRDNTDEMILVNNKHAIYKKNATSL--SLRRDIRNAPNKKI-TLDDTTAIKSTSS 457

QY 572 ELARSNTVDLSALLNPSETVQTRGVSFVPSNOHRRSRSTIEGGIETVNNASLLKTSS 631

QY 458 VQFAMLOFLYDHIQTHINDFSRIATAWCELQNRVLEWHGKINFSASATLGRVA 517

QY 632 VEFAMIQFAYDYIAHVNEMLSRATATWCTIQNRHVLWTETLKLNGGVVSMALERSV 691

QY 518 AKMLGDVAVSSCTAIDAESVYTLQNSMRVTITSTNWCYSRPLVLSYGENOGTQQLGEN 577

QY 692 ARLLGDVAVTQCVNISSGHVYIQNSMRVTGSSSTCYSRPLVSFRALNDSEYIEGQLGEN 751

QY 578 NELLPTLEAVEPCSNHRRYFLEGGYALPENYFVKMVDAAQIOIASTFVELNLTILED 637

QY 752 NDLIVERKLEPCTVNNKRYKFGADYVFEYAYVRKVLPELSEIELISAYVDLNLTTLED 811

QY 638 REILPLSVYKTELDRGVLDYAEVARRNOLHELK 672

QY 812 REFLEPVYTRAELEDGLDYSEIQNRNQLHALK 846

RESULT 11
US-08-413-118-4
Sequence 4, Application US/08413118
Patent No. 5688920
GENERAL INFORMATION:
APPLICANT: PAOLETTI, ENZO
APPLICANT: LIMBACH, KEITH J..
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.

STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413.118
FILING DATE: 29-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220.151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-413-118-4

Query Match 50.9%; Score 1786.5; DB 1; Length 1041;
Best Local Similarity 45.4%; Pred. No. 8.3e-156;
Matches 343; Conservative 126; Mismatches 185; Indels 101; Gaps 5;

QY 18 GTNSSPSTQNTVTSREVSSVOLSEEE-----STFYLCPPPVGTVIRLEPPKCPPEPRKAT 73
DB 93 GTPPKPTDPTDMSDREALRASQIEANGSTFTYMCPPSGTVVRLPPRACPDYKLGK 152

QY 74 EWGECIALFKENISPVYKFKVLYKNIIQTTTGTGTYRQITNRYTDRTPVSEIEITDL 133
DB 153 NFTGIAVIFKENIAPYKFKANIYKNIIMTVWSSGYATVNNRYTDRVPKVOEITDL 212

QY 134 IDGGRSSKARYLRNNVYVEAFDRDAGEKQVLKPSKFNTPESR----- 178
DB 213 IDRGMCILSKADYVRNNYQFTAFDRDEDPRELPLKPSKFNTPQSRGWHYKFKATVYVKD 272

QY 179 ----- 178

QY 273 VIVSTAWAGSSYQITNRYADRPVSEITDIDKFGKCSKATYVRNNHKEAFNEDEK 332

QY 179 -----AWHTTNETVTVMGSPWYRTGTSVNCIVEEMDARSVPYYSF 220

QY 333 NPQDMPLIASKYNSVSGKAWHTTNETYTKGAAGFHSGTSVNCIVEEDARSVPYDSF 392

QY 221 AMANGDANTISPFYGLSPPEAAAPMGYPQDNFKQSDYSFMDLDRKRSKASLPYKRNFLI 280

QY 393 AISTGDVIHMSPPFGLR-DGAHVEHTSYSSDRFQIQIEGYPIDLDTLRLQLGAPVSRNLE 451

QY 281 TSHTFTGWDWAPKTRVCSMTKWEVTEMLRATVNGRYRPMARELSATFISNTTFEPDNR 340

QY 452 TPHYVAVNWTPKCGRVCTIAKWEIDEMLRDEYQGSYRFTVTKTISATFISNTSQFEINR 511

QY 341 IILGQCICKREAAAEIQRFTKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLR 400

QY 512 IRLGDCATKEAAEAIDRIYKSKYKTHIQGTLETYLARGGFLIAPRPMISNELAKLYIN 571

QY 401 ELMRDNTDEMILVNNKHAIYKKNATSL--SLRRDIRNAPNKKI-TLDDTTAIKSTSS 457

QY 572 ELARSNTVDLSALLNPSETVQTRGVSFVPSNOHRRSRSTIEGGIETVNNASLLKTSS 631


```

:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Frommer, William S.
: REGISTRATION NUMBER: 25,506
: REFERENCE/DOCKET NUMBER: 454310-2540
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 840-3333
: TELEFAX: (212) 840-0712
: TELEX: 425066 CURTMS
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 868 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: N-terminal
: US-08-220-151-7

Query Match 49.2%; Score 1726; DB 1: Length 868;
Best Local Similarity 50.5%; Pred. No. 2.4e-150;
Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;

QY 31 REVSSVOLSEESTFVCLPPPGSVIRLEPPKCEPRKATWEGBIAILFKENISPY 90
DB 41 REAHKHSQDAETPEFTYVCPPTGSTRLEPTPCDYHLGKNFTGIAVYVKENIAAY 100
QY 91 KFKVLYYKNIOTTTWGTYYQITNRYTDRTPVSLIEELTDLIDGKRCSSKARYLRNN 150
DB 101 KFKATVYKDVYSTAWAGSSYQITNRYADRPVPIVSEITDITDKFGKCSSKATYYVRNN 160
QY 151 VYVEAFDRDAGEKOVLLKPSKFNTPESRAHHTTNETYTVMGSPWIYRTGTSVNCIVBEMD 210
DB 161 HKVEAFNEEDKNPDMLPIASKYNSVGSKAHHTNDTYMVAGTPTGYRTGTSVNCIIIEVE 220
QY 211 ARSVEPYSVFAMANGDIANISPEYGLSPPEAAAEPMGYPODNFKQLDYSYFMDLDRKKA 270
DB 221 ARSIFPDSELTGDIITYNSPEFGLR-DGAYREHSNYAMDRPHQFEGYRQRLDTRALL 279
QY 271 SLPVKNFLTSHFTVGWDAWAPTTRVCSMTKWKVETEMLRATVNGRYRPMARELSATFI 330
DB 280 E-PAARNFLVTPHLTVGNWKKPKRTEVCSLVKKREVEDVYRDEYAHNFRFTMKLTSTFTFI 338
QY 331 SNITEPDNRIILGQCIKREAAEAIQIFRTKYNDSHVKGHVQYFLALGGFTIVAYQPVIL 390
DB 339 SETNEFNLOIHLSQCVCBEARAIIRIITRYNSSHVRTGDIQTYLARGGFVVFQPLL 398
QY 391 SKSLAHMYRELNRDNRITDEMLDLVNKKAIYKKNATSLSLRLARDLRNAPNRKITLDDTT 450
DB 399 SNSLARLYLQELVRE-----WNHS-PQKHPTRWTRSRSV-----PVELRANR 441
QY 451 ATKSTSVQFAMQLFYDHIQTHINDMFRIATAWCELQNRELVLHHEGIKINPSATASA 510
DB 442 TITTTSSVFAMQLFYDHIQEVNEMLARISSSWCQLQNRERALSGLFPIINPSALAST 501
QY 511 TLGRRVAAKMLGDVAASCTAIDAES-VTLQNSMEVITSTNFCYSRPLVLFSGYGENQGN 569
DB 502 ILDQVKARILGDVIVSNCPDELGSDTRILLQNSMRVSGSTTRCYSRPLISIVSLNGSGT 561
QY 570 IQQGLGENNELPTLEAVPCPSANHRRYFLFGSGYALFENYFNKVMVDADIAQIASTFVE 629
DB 562 VEGQLGTDELIMSRLLEPCVANHKRYFLFGHHVYVYEDYRVYREIAVHDVGMISTYVD 621
QY 630 LNLITLEDREILPLSYTTEELRDGVCLDYAEVARNRQJHELK 672
DB 622 LNLTLTKDREFPLQVYTRDELDTGLLDYSIQIRNONMHSR 664

```

RESULT 14
US-08-413-118-7
; Sequence 7, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:

APPLICANT: PAOLETTI, ENZO
 APPLICANT: LIMBACH, KEITH J.
 TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 NUMBER OF SEQUENCES: 128
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 STREET: 530 FIFTH AVENUE, 25TH FLOOR
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/413,118
 FILING DATE: 29-MAR-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/220,151
 FILING DATE: 30-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: FROMMER, WILLIAM S.
 REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454310-2670
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 868 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: N-terminal
 US-08-413-118-7

Query Match	49.24;	Score 1726;	DB 1;	Length 868;
Best Local Similarity	50.5%;	Pred. No. 2.4e-150;		
Matches	325;	Conservative 116;	Mismatches 182;	Indels 20;
Gaps	6			
QY	31	REVSSVOLSEESTFVLCPPGVSTYRLPPKCPPEPRATEWGEIGAILFKENISPY	90	
Db	41	REALHKSDAETKFTFVCPPTGSTIVRLPPTCTPDYHLGKNFTGEIVVYKENIAY	100	
QY	91	KFKVTLVYKNIQTTTGTGYRTNRYTDTVPVSEIEITDLIDGKRCSSKARYLNN	150	
Db	101	KFKATVYKDVIVSTAWAGSYTQITNRYADRPVPVSEITDITDKFKCSKATYVRN	160	
QY	151	VYVEAFDRAGEROVLLKPSKFNTPESRAWHTTNETYVWGSPWLYRTGTSVNCIVE	210	
Db	161	HKVEAFNEDKNQDMPITASIKSVGKAWHITNDTYNVAGTPTGYRTGTSVNCII	220	
QY	211	ARSVPFYSFAMANGDIANSPPFYGLSPPEAAAPMGYPQNFQOLDYSFSDLDKRKA	270	
Db	221	ARSIFFYDSFGLSTGDIYHNSPPFGLR-DGAYREHSNVAMDRFHQFEGYRQDLD	279	
QY	271	SLPVKRNFLTISHTYQWCDWNAKPTRVCSMTKWKEVTEMLRATVNGRYRFMA	330	
Db	280	E-PAARNFLVTPHLTVGWNKKPKRTECVSLVKWEVEDVVRDEYAHNFRFTMKLT	338	
QY	331	SNTEFDPNRIILGQCILKREAAAEIQIFRTKYNDSHVKVGHVOYFLALGCFIVAY	390	
Db	339	SETNEFNLIQHLSQCVKKEARAIINRIYTRYNSSHVRTGDIQTYLARGFVVVF	398	
QY	391	SKSLAHMYLRELMNRNRTDEMLDVNNKHAIFYKKNAATSLSLRRDRIRNAPNR	450	
Db	399	SNSLRIYLOELVRE-----NPNHS-PQKHPTNRTSRSRV-----PVELRANR	441	

Db	339	SETNEFNLIHQISQCVCKEARALINRIYTRYNSSHVRTGDIQTLYLARGGVVVFQPLL	3398
QY	391	SKSLAHMYLRELMDRNTDEMLDVNNKHAIYKKNATSLSLRDLRINAPNRKTTDDTT	450
Db	399	SNSLIARIYLQELVRE-----NTNHS-PQKHPTNTRSRSV-----PVELRANR	441

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:58:45 ; Search time 93.4809 Seconds
(without alignments)
2997.887 Million cell updates/sec

Title: US-09-147-052-4
Perfect score: 5619
Sequence: 1 MHYFRNCIFFLVILYGTN.....SSNNADKIPGVRPCTFL 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL_23:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3598	64.0	865	12 Q8JLW4	Q8JLW4 turkey herp
2	3595.5	64.0	805	12 Q98Y44	Q98Y44 turkey herp
3	3595.5	64.0	865	12 Q8JLW2	Q8JLW2 turkey herp
4	3588.5	63.9	805	12 Q98Y45	Q98Y45 turkey herp
5	3587	63.8	865	12 Q8JLW3	Q8JLW3 turkey herp
6	3585	63.8	865	12 Q83291	Q83291 turkey herp
7	3089	55.0	865	12 Q9PW21	Q9PW21 turkey herp
8	3074	54.7	865	12 Q69406	Q69406 turkey herp
9	3054.5	54.4	870	12 Q9DPQ9	Q9DPQ9 meleagrid h
10	3048.5	54.3	864	12 Q69408	Q69408 meleagrid h
11	3048.5	54.3	870	12 Q9EIG4	Q9EIG4 meleagrid h
12	1965	35.0	879	12 Q91SD0	Q91SD0 canine herp
13	1930	34.3	881	12 Q84735	Q84735 phocine herp
14	1909.5	34.0	948	12 Q90050	Q90050 feline herp
15	1805	32.1	919	12 Q9QAP5	Q9QAP5 caprine herp
16	1803	32.1	931	12 Q9QAP6	Q9QAP6 rangiferine

17	1783.5	31.7	868	12 Q98VM4	Q98VM4 human herp
18	1779.5	31.7	980	12 Q66882	Q66882 equine herp
19	1763.5	31.4	950	12 Q66678	Q66678 equine herp
20	1756	31.3	906	12 Q91WV3	Q91WV3 pseudorabie
21	1755.5	31.2	975	12 Q39275	Q39275 equine herp
22	1747.5	31.1	891	12 Q9J052	Q9J052 simian herp
23	1747.5	31.1	929	12 Q9QAP4	Q9QAP4 cervid herp
24	1747	31.1	947	12 Q9QAP7	Q9QAP7 bovine herp
25	1746.5	31.1	912	12 Q8QV01	Q8QV01 buffalo herp
26	1743	31.0	943	12 Q86665	Q86665 feline herp
27	1740.5	31.0	891	12 Q66018	Q66018 simian herp
28	1739	30.9	904	12 Q91C63	Q91C63 herpes simp
29	1738	30.9	904	12 Q89920	Q89920 herpes simp
30	1737	30.9	901	12 Q94465	Q94465 herpes simp
31	1735	30.9	904	12 Q37453	Q37453 herpes simp
32	1734	30.9	901	12 Q69464	Q69464 herpes simp
33	1729.5	30.8	920	12 Q8JY98	Q8JY98 herpesvirus
34	1721	30.6	902	12 Q69095	Q69095 herpes simp
35	1719.5	30.6	885	12 Q69387	Q69387 cercopithec
36	1712.5	30.5	908	12 Q65338	Q65338 baboon herp
37	1707.5	30.4	893	12 Q65540	Q65540 baboon herp
38	1706.5	30.4	904	12 Q69526	Q69526 human herp
39	1705.5	30.4	904	12 Q9QLM8	Q9QLM8 human herp
40	1705	30.3	916	12 Q66016	Q66016 cercopithec
41	1702.5	30.3	903	12 Q69076	Q69076 human herp
42	1702.5	30.3	904	12 Q91WU4	Q91WU4 human herp
43	1698.5	30.2	904	12 Q9DXE4	Q9DXE4 herpes simp
44	1667	29.7	894	12 Q9YW92	Q9YW92 macropodid
45	1653	29.4	887	12 Q9YW93	Q9YW93 macropodid

ALIGNMENTS

RESULT 1

Q8JLW4 ID Q8JLW4 PRELIMINARY; PRT; 865 AA.

AC Q8JLW4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.

OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_taxID=10390;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-X;
RA Parcells M.S., Shamblin C.E., Dienglewiecz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (VMDV, vvMDV, vv+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129965; AAM97698.1; -;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98107 MW; 4D2628B5E4DEB28B CRC64;

Query Match 64.0%; Score 3598; DB 12; Length 865;
Best Local Similarity 85.8%; Pred. No. 4.6e-163;
Matches 718; Conservative 17; Mismatches 60; Indels 42; Gaps 8;

QY 1 MHYFRNCIFFLVILYGTNSSPSTQNTVSREVSSVQLSEESTFFILCPVGVSTVIRL 60
Db 1 MHYFRNCIFFLVILYGTNSSPSTQNTVSREVSSVQLSEESTFFILCPVGVSTVIRL 60
QY 61 EPPKRCPEPRKATEWEGGIALFKENISPKYFKVTLVYKNIQTFTWTGTYRITNRYT 120
Db 61 EPPKRCPEPRKATEWEGGIALFKENISPKYFKVTLVYKNIQTFTWTGTYRITNRYT 120

```

QY 121 DRTPVSEIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPPESRAW 180
Db 121 DRTPVSEIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPPESRAW 180

QY 181 HTTNETYTVWGSPIIYRTGTSYNCIIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPE 240
Db 181 HTTNETYTVWGSPIIYRTGTSYNCIIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPE 240

QY 241 AAAPMGYPQDNFKQLDYSFSDMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTTTRVCSM 300
Db 241 AAAPMGYPQDNFKQLDYSFSDMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTTTRVCSM 300

QY 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360

QY 361 TYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYLRELNRDRTDEMDLVNNKHA 420
Db 361 TYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYLRELNRDRTDEMDLVNNKHA 420

QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAITSSTSSVQFAMQLQFLYDHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAITSSTSSVQFAMQLQFLYDHIQTHINDMFSR 480

QY 481 IATAWCELONRELVLWHGEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
Db 481 IATAWCELONRELVLWHGEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540

QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQGNIOGOLGENNELPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFYSGENQGNIOGOLGENNELPTLEAVEPCSANHRRYFLF 600

QY 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDRILPLSVYTKKEELRDVGVLDA 660
Db 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDRILPLSVYTKKEELRDVGVLDA 660

QY 661 EVARRNQLHELKFDYDINKVIEVDNTYA---GLOEFGCMSITKKDANPNNGQTOL-EAARM 716
Db 661 EVARRNQLHELKFDYDINKVIEVDNTYA---GLOEFGCMSITKKDANPNNGQTOL-EAARM 716

QY 717 ELTDLINAKAMTASLQDYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758
Db 717 ELTDLINAKAMTASLQDYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758

QY 759 -TLEOLKMAKTNLESAINQANTDKTTFD-----YKAKTTTLEQATNL 809
Db 759 -TLEOLKMAKTNLESAINQANTDKTTFD-----YKAKTTTLEQATNL 809

QY 769 MTEVLKKAQATRELHGEESDDLERTSIDER--KLEAREMIKYMALVSAERHEKKL 823
Db 769 MTEVLKKAQATRELHGEESDDLERTSIDER--KLEAREMIKYMALVSAERHEKKL 823

RESULT 2
Q98Y44
ID Q98Y44 PRELIMINARY; PRT; 805 AA.
AC Q98Y44;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Glycoprotein B (Fragment).
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson I.;
RT "Marek's disease virus - turkey isolate, gB.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350326; AAK37557.2;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
FT NON_TER 805
SQ SEQUENCE 805 AA; 91115 MW; FC8EB090885E15AA CRC64;

```

```

Query Match 64.0%; Score 3595.5; DB 12; Length 805;
Best Local Similarity 88.0%; Pred. No. 5.5e-163;
Matches 711; Conservative 14; Mismatches 48; Indels 35; Gaps 6;

QY 1 MHYFRNCIFFLIVILYGTNSSPSQNTVTSREVWSSVQLSEESTFYLCPPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVILYGTNSSPSQNTVTSREVWSSVQLSEESTFYLCPPPVGVSTVIRL 60

QY 61 EPPKCPPEPRKATEMGEIGAILFKENISPYKFKVLYYKNIQTITTTGTCTTQIINRYT 120
Db 61 EPPKCPPEPRKATEMGEIGAILFKENISPYKFKVLYYKNIQTITTTGTCTTQIINRYT 120

QY 121 DRTPVSEIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPPESRAW 180
Db 121 DRTPVSEIEITDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPPESRAW 180

QY 181 HTTNETYTVWGSPIIYRTGTSYNCIIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPE 240
Db 181 HTTNETYTVWGSPIIYRTGTSYNCIIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPE 240

QY 241 AAAPMGYPQDNFKQLDYSFSDMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTTTRVCSM 300
Db 241 AAAPMGYPQDNFKQLDYSFSDMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTTTRVCSM 300

QY 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFR 360

QY 361 TYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYLRELNRDRTDEMDLVNNKHA 420
Db 361 TYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYLRELNRDRTDEMDLVNNKHA 420

QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAITSSTSSVQFAMQLQFLYDHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAITSSTSSVQFAMQLQFLYDHIQTHINDMFSR 480

QY 481 IATAWCELONRELVLWHGEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540
Db 481 IATAWCELONRELVLWHGEGIKINPSATASATLGRVAAKMLGDVAASVSCTAIDAESVTL 540

QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQGNIOGOLGENNELPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFYSGENQGNIOGOLGENNELPTLEAVEPCSANHRRYFLF 600

QY 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDRILPLSVYTKKEELRDVGVLDA 660
Db 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDRILPLSVYTKKEELRDVGVLDA 660

QY 661 EVARRNQLHELKFDYDINKVIEVDNTYA---GLOEFGCMSITKKDANPNNGQTOL-EAARM 716
Db 661 EVARRNQLHELKFDYDINKVIEVDNTYA---GLOEFGCMSITKKDANPNNGQTOL-EAARM 716

QY 717 ELTDLINAKAMTASLQDYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758
Db 717 ELTDLINAKAMTASLQDYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758

QY 709 VVVGGAAGAIYSTISGSFASFPSCALGIIAGLVAATLAYRYVKNLKSNNPKALYP 768
Db 709 VVVGGAAGAIYSTISGSFASFPSCALGIIAGLVAATLAYRYVKNLKSNNPKALYP 768

QY 759 -TLEOLKMAKTNLESAINQANTDKTTFD 785
Db 759 -TLEOLKMAKTNLESAINQANTDKTTFD 785

QY 769 MTEVLKKAQATRELHGEESDDLERTSID 796
Db 769 MTEVLKKAQATRELHGEESDDLERTSID 796

RESULT 3
Q9JLW2
ID Q9JLW2 PRELIMINARY; PRT; 865 AA.
AC Q9JLW2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.

```


OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=attenuated GA;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (vMDV, vMDV, vMDV, vMDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vv+MDVs."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129969; AA097702.1; -
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98020 MW; F4F526114EA02DE5 CRC64;

Query Match 64.0%; Score 3595.5; DB 12; Length 865;
Best Local Similarity 84.1%; Pred. No. 6e-163;
Matches 723; Conservative 18; Mismatches 68; Indels 51; Gaps 9;

QY 1 MHYFRNCIFFLLVILYGNSSPSTQNTVSREVSSVQLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLLVILYGNSSPSTQNTVSREVSSVQLSEESTFYLCPPPGVSTVIRL 60

QY 61 EPRKCPERKATWEGEGIAILFKENISPKFKVTLVYKNIQTTTWTGTYQITNRYT 120
DB 61 EPRKCPERKATWEGEGIAILFKENISPKFKVTLVYKNIQTTTWTGTYQITNRYT 120

QY 121 DRTPVSEIETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPSRAW 180
DB 121 DRTPVSEIETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPSRAW 180

QY 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240

QY 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240

QY 241 AAAEPNGYQDQNFQKLDVSFMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300
DB 241 AAAEPNGYQDQNFQKLDVSFMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300

QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTDPNRIILGQCIRKREAAIEQIFR 360
DB 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTDPNRIILGQCIRKREAAIEQIFR 360

QY 361 TKYNDSHVKVGHVQYFLAAGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLYNKKHA 420
DB 361 TKYNDSHVKVGHVQYFLAAGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLYNKKHA 420

QY 421 IYKKNATSLRLRRDRNAPNRKITLDDTTAISTSSVOFAMQLQFLYDHIQTHINDMFSR 480
DB 421 IYKKNATSLRLRRDRNAPNRKITLDDTTAISTSSVOFAMQLQFLYDHIQTHINDMFSR 480

QY 481 IATAWCELQNLRELVLWHEGIIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAEVTL 540
DB 481 IATAWCELQNLRELVLWHEGIIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAEVTL 540

QY 541 QNSMRVITSTNTCYSPRLVFSYGENQNIQOLGENNELLPLEAVEPCSANHRYFLF 600
DB 541 QNSMRVITSTNTCYSPRLVFSYGENQNIQOLGENNELLPLEAVEPCSANHRYFLF 600

QY 601 GSGYALFENNFVKWDAADIQIATFVFNLTLLDRETLPLSVYTKELRDVGVLDA 660
DB 601 GSGYALFENNFVKWDAADIQIATFVFNLTLLDRETLPLSVYTKELRDVGVLDA 660

QY 661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQFCGMSITRKDANPNNGOTOL-BAARM 716
DB 661 EVARRNQLHELKPYDINKVIEVDNTYAFNGLAEL-----FNGMGVQVGAIGK 708

QY 717 ELTDLINAKAMTLASLDYAK-----IEASLSAYSAAETVN---NNLNA--- 758
DB 709 VVVGAGAIVSTISGVSFAFMSNPFAGALIGLIIAGLVAFLAYRYVKNLKSMPKALVP 768
QY 759 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEA-----YKAL-----KYLE 803
DB 759 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEA-----YKAL-----KYLE 803

DB 769 MTTEVLKKAQATRELHGEESDLDERTSIDER--KLEAREMIKYMALVSAERHEKKLRRE 836
QY 804 QRATNLEGLSSTAYNQIRNN 823
DB 827 RRGCTTAVLSHDLAKMRKNS 846

RESULT 4
Q98Y45 PRELIMINARY; PRT; 805 AA.
ID Q98Y45;
AC Q98Y45;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B (Fragment).
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson I.;
RT "Marek's disease virus - chicken isolate, gb gene."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350325; AA037556.2; -
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
FT NON_TER 805
SQ SEQUENCE 805 AA; 91111 MW; F88B4410A85E15AA CRC64;

Query Match 63.9%; Score 3588.5; DB 12; Length 805;
Best Local Similarity 87.9%; Pred. No. 1.2e-162;
Matches 710; Conservative 14; Mismatches 49; Indels 35; Gaps 6;

QY 1 MHYFRNCIFFLLVILYGNSSPSTQNTVSREVSSVQLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLLVILYGNSSPSTQNTVSREVSSVQLSEESTFYLCPPPGVSTVIRL 60

QY 61 EPRKCPERKATWEGEGIAILFKENISPKFKVTLVYKNIQTTTWTGTYQITNRYT 120
DB 61 EPRKCPERKATWEGEGIAILFKENISPKFKVTLVYKNIQTTTWTGTYQITNRYT 120

QY 121 DRTPVSEIETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPSRAW 180
DB 121 DRTPVSEIETDLDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPSRAW 180

QY 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVGSPWIIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISPFYGLSPPE 240

QY 241 AAAEPNGYQDQNFQKLDVSFMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300
DB 241 AAAEPNGYQDQNFQKLDVSFMDLDRKRRKASLPVKRNFLITSHFTVGDWAPKTRVCSM 300

QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTDPNRIILGQCIRKREAAIEQIFR 360
DB 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTDPNRIILGQCIRKREAAIEQIFR 360

QY 361 TKYNDSHVKVGHVQYFLAAGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLYNKKHA 420
DB 361 TKYNDSHVKVGHVQYFLAAGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLYNKKHA 420

QY 421 IYKKNATSLRLRRDRNAPNRKITLDDTTAISTSSVOFAMQLQFLYDHIQTHINDMFSR 480
DB 421 IYKKNATSLRLRRDRNAPNRKITLDDTTAISTSSVOFAMQLQFLYDHIQTHINDMFSR 480

QY 481 IATAWCELQNLRELVLWHEGIIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAEVTL 540
DB 481 IATAWCELQNLRELVLWHEGIIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAEVTL 540

QY 541 QNSMRVITSTNTCYSPRLVFSYGENQNIQOLGENNELLPLEAVEPCSANHRYFLF 600
DB 541 QNSMRVITSTNTCYSPRLVFSYGENQNIQOLGENNELLPLEAVEPCSANHRYFLF 600

```

Db 541 QNSMRVITSTNTCYSRPLVLSYGENQNIQGLGNNELLPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKWDAADIOIASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
Db 601 GSGYALFENYFVKWDAADIOIASTFVELNLTLEDRILPLSVYTKELRDVGVLDA 660
QY 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFGCMSITKKDANPNNGOTQL-EAARM 716
Db 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFGCMSITKKDANPNNGOTQL-EAARM 716
QY 717 ELTDLINAKAMTSLASQDYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758
Db 709 VVGAAGAIVTISGVSAFMSNPFGLAIGLIIAGLVAFLAYRVNKLKSNPKALYP 768
QY 759 -TLEOLKMAKTNLESAINQANTDKTTFD 785
Db 769 MTEVLKQAQATRELHGEESDDLERTSID 796

RESULT 5
Q8JLW3
ID Q8JLW3 PRELIMINARY; PRT; 865 AA.
AC Q8JLW3;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N;
RA Parcellis M.S., Shamblin C.E., Dienglewiez R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (VMDV, vvMDV, vv+MDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY129967; AA07700.1; -
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98030 MW; 9008B58B4E114D CRC64;

Query Match 63.8%; Score 3587; DB 12; Length 865;
Best Local Similarity 85.7%; Pred. No. 1.5e-162;
Matches 717; Conservative 17; Mismatches 61; Indels 42; Gaps 8;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSSVQLSEESTFYLCPPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSSVQLSEESTFYLCPPPVGVSTVIRL 60
QY 61 EPPKCEPKRATKATWEGGIALFKENISPYKFKVLYYKNIQTTWTGTYRQITNRYT 120
Db 61 EPPKCEPKRATKATWEGGIALFKENISPYKFKVLYYKNIQTTWTGTYRQITNRYT 120
QY 121 DRTPVSTEEITDLDGRCSSKARYLRNNVYVEAFDRDAGEQVLLKPSKFTNPESRAW 180
Db 121 DRTPVSTEEITDLDGRCSSKARYLRNNVYVEAFDRDAGEQVLLKPSKFTNPESRAW 180
QY 181 HTTNETYVVGSPWYRTGTQSVNCIVEEMDARSVPYSYFAMANGDIANSPPYGLSPPE 240
Db 181 HTTNETYVVGSPWYRTGTQSVNCIVEEMDARSVPYSYFAMANGDIANSPPYGLSPPE 240
QY 241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVKRNFLTISHTFGWDWAPKTTTRVCSM 300
Db 241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVKRNFLTISHTFGWDWAPKTTTRVCSM 300
QY 301 TKKKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQICKREAAIAEQIFR 360
Db 301 TKKKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILGQICKREAAIAEQIFR 360

```

RESULT 6

```

Q83291
ID Q83291 PRELIMINARY; PRT; 865 AA.
AC Q83291;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Glycoprotein 100 precursor.
GN GB.
OS Marek disease virus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=38013;
RN [1]
RP SEQUENCE OF 1-498 FROM N.A.
RC STRAIN=JM;
RA Sousloparov M.A., Bakhtina M.M., Krendelshtchikov A.V., Babkin I.V.;
RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
RT complex gp100, gp60, gp48 (b-antigen).";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X91985; CAA63039.1; -
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;

Query Match 63.8%; Score 3585; DB 12; Length 865;
Best Local Similarity 85.7%; Pred. No. 1.9e-162;
Matches 717; Conservative 16; Mismatches 62; Indels 42; Gaps 8;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSSVQLSEESTFYLCPPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSSVQLSEESTFYLCPPPVGVSTVIRL 60
QY 61 EPPKCEPKRATKATWEGGIALFKENISPYKFKVLYYKNIQTTWTGTYRQITNRYT 120
Db 61 EPPKCEPKRATKATWEGGIALFKENISPYKFKVLYYKNIQTTWTGTYRQITNRYT 120

```

```

QY 121 DRTPTVSEITDLIDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAW 180
DB 121 DRTPTVSEITDLIDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAW 180
QY 181 HTTNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGLSPPE 240
DB 181 HTTNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGLSPPE 240
QY 241 AAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCVSM 300
DB 241 AAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCVSM 300
QY 301 TWKKEVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCIREAAAEIQIFR 360
DB 301 TWKKEVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCIREAAAEIQIFR 360
QY 361 TYKNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYRLMRDNRDTEMLDLVNKKH 420
DB 361 TYKNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYRLMRDNRDTEMLDLVNKKH 420
QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHIQTHINDMFSR 480
DB 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHIQTHINDMFSR 480
QY 481 IATANCELORELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAESVTL 540
DB 481 IATANCELORELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAESVTL 540
QY 541 QNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
DB 541 QNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLEDREILPLSVYTKELRDVGVLDA 660
DB 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLEDREILPLSVYTKELRDVGVLDA 660
QY 661 EVARRNQLHELFFYDINKVIEDTNYA---GLQFGCMSITKNDANPNNGOTQL-EAARM 716
DB 661 EVARRNQLHELFFYDINKVIEDTNYA---GLQFGCMSITKNDANPNNGOTQL-EAARM 716
QY 717 ELTDLINAKMTLASLDQYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758
DB 717 ELTDLINAKMTLASLDQYAK-----IEASLSAYSEAEVTN-----NNLNA--- 758
QY 709 VVVGAGAIYVTSIGVSAPMSNPFNALIGLIIAGLVAFLAYRYVKNLKNPKALYP 768
DB 709 VVVGAGAIYVTSIGVSAPMSNPFNALIGLIIAGLVAFLAYRYVKNLKNPKALYP 768
QY 759 -TLEQKMAKTNLESAINQANTDKTTFDNEHNPVLEA-----YKAKTTTLEQATNL 809
DB 759 -TLEQKMAKTNLESAINQANTDKTTFDNEHNPVLEA-----YKAKTTTLEQATNL 809
QY 769 MTEVLKAQATRELHGEESDDLERTSIDER--KLEAREMIKYMALVSAEERHEK 823
DB 769 MTEVLKAQATRELHGEESDDLERTSIDER--KLEAREMIKYMALVSAEERHEK 823

RESULT 7
Q9PWZ1 ID Q9PWZ1 PRELIMINARY; PRT; 865 AA.
AC 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
DE Homolog of HSV-1 major DNA-binding protein (Glycoprotein B).
GN ORF 35 OR UL27.
OS Turkey herpesvirus,
OS Marek's disease virus serotype 2 MDV2, and
OS Gallid herpesvirus 3.
OC Viruses; gsdna viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390, 36353, 35250;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Turkey herpesvirus; STRAIN-HPRS24;
RA Kato K., Jang H., Izumiya Y., Cai J., Tsushima Y., Miyazawa T.,
RA Kai C., Mikami T.;
RT "Identification and Transcriptional Analysis of the Marek's Disease
RT Virus Serotype 2 Genes Homologous to the Glycoprotein B (UL27), the
RT ICP18.5 (UL28) and the Major DNA-binding Protein (UL29) Genes of

```

```

RT Herpes Simplex Virus Type 1.;
RL J. Vet. Med. Sci. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 1; STRAIN-HPRS24;
RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
RA Lee Y., Kai C., Takahashi E., Mikami T.;
RT "The complete DNA sequence and transcription map of the unique long
RT genome region of Marek's disease virus type 2.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 3; STRAIN-HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-Gallid herpesvirus 3; STRAIN-HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RT "A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
RT Strain HPRS24.";
RL Curr. Top. Microbiol. Immunol. 0:0-0(2000).
DR EMBL; AB024711; BA83751.1; -.
DR EMBL; AB024414; BA82923.1; -.
DR EMBL; AB049735; BAB16537.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW DNA-binding.
SQ SEQUENCE 865 AA; 97309 MW; A20B66C70A475C35 CRC64;

Query Match 55.0%; Score 3089; DB 12; Length 865;
Best Local Similarity 70.4%; Pred. No. 7e-139;
Matches 613; Conservative 80; Mismatches 134; Indels 44; Gaps 10;

QY 1 MHYFRNCI-FELIVLYGTNSPSTQNTSREWSSVOLSEESFYLCPPVGVGTIVR 59
DB 1 MHYFRNCI-FELIVLYGTNSPSTQNTSREWSSVOLSEESFYLCPPVGVGTIVR 59
QY 60 LPPPRKCPKPRKATGEGIAILFKENISPYKPKVLYYKNIQTTTGTGTYRQITNRY 119
DB 60 LPPPRKCPKPRKATGEGIAILFKENISPYKPKVLYYKNIQTTTGTGTYRQITNRY 119
QY 61 LEPPRCPEPLKATGEGIAILFKENINPYKFKVLYYKNIQTTTGTGTYRQITNRF 120
DB 61 LEPPRCPEPLKATGEGIAILFKENINPYKFKVLYYKNIQTTTGTGTYRQITNRF 120
QY 120 TDRTPVSEITDLIDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRA 179
DB 120 TDRTPVSEITDLIDGKRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRA 179
QY 121 TDRTPVSEITDLIDGKRCSSKARYLRNNVVEAYDGDAGEKQVLLKPSKENTPESKA 180
DB 121 TDRTPVSEITDLIDGKRCSSKARYLRNNVVEAYDGDAGEKQVLLKPSKENTPESKA 180
QY 180 WHITNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGLSP 239
DB 180 WHITNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGLSP 239
QY 181 WHITNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGVAPP 240
DB 181 WHITNETTYTWGSPWYRTGTSVNCIVEEMDARSVPYSYFANANGDIANISFYGVAPP 240
QY 240 EAAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCV 299
DB 240 EAAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCV 299
QY 241 EAAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCV 300
DB 241 EAAAEPMGYQDQNFKQLDYSFMDLKRKASLPVKNFLITSHFTVGVWDWAPKTTVCV 300
QY 300 MTKWKEVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCIREAAAEIQIF 359
DB 300 MTKWKEVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCIREAAAEIQIF 359
QY 301 MVKWKDVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCVRDAETTKIRF 360
DB 301 MVKWKDVTEMLRATVNGRYRFMARELSATFISNTTFDPNRIILGQCVRDAETTKIRF 360
QY 360 RTKYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYRLMRDNRDTEMLDLVNKKH 419
DB 360 RTKYNDSHVKGHVQVFLALGGFIVAYQPVLSKSLAHMYRLMRDNRDTEMLDLVNKKH 419
QY 361 AOKYNDSHVKGHVQVFLALGGFIVAYQPVMSKSLAHMYRLMRDNRDTEMLDLVNKKH 420
DB 361 AOKYNDSHVKGHVQVFLALGGFIVAYQPVMSKSLAHMYRLMRDNRDTEMLDLVNKKH 420
QY 420 AIYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHIQTHINDMFS 479
DB 420 AIYKKNATSLRLRRDIRNAPNRKITLDDTTAISTSSVQFAMQLFLYDHIQTHINDMFS 479
QY 421 ALSNKTNTVSLRLRELNRNAPKEKALNEGATIRSTSSVQFAMQLFLYDHIQTHINDMFS 480
DB 421 ALSNKTNTVSLRLRELNRNAPKEKALNEGATIRSTSSVQFAMQLFLYDHIQTHINDMFS 480
QY 480 RIATANCELORELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAESVT 539
DB 480 RIATANCELORELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSCCTAIDAESVT 539
QY 481 RIATANCELORELVLWHEGIKINPSA IASATLGRVAAKMLGDVTA VSTCTGIDAESVT 540
DB 481 RIATANCELORELVLWHEGIKINPSA IASATLGRVAAKMLGDVTA VSTCTGIDAESVT 540
QY 540 LQNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFL 599
DB 540 LQNSMRVITSTNTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFL 599

```


Best Local Similarity 68.2%; Pred. No. 3.1e-137;		Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	
Matches 602; Conservative 88; Mismatches 132; Indels 61; Gaps 11;		Alphanepesvirinae; Marek's disease-like viruses.	
QY	1 MHYFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVOLSEESTFYLCPPPVGVSTVRL 60	OC	1 MHYFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVOLSEESTFYLCPPPVGVSTVRL 60
Db	7 MKYFNLSFLFPLPILLSIATSEIKLPNTVAREIVSGIQLSEDETFYVCPVPVGSIVRL 66	OX	7 MKYFNLSFLFPLPILLSIATSEIKLPNTVAREIVSGIQLSEDETFYVCPVPVGSIVRL 66
QY	61 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 120	RN	61 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 120
Db	67 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 126	RP	67 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 126
QY	121 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 180	RC	121 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 180
Db	127 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 186	RX	127 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 186
QY	181 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 240	RA	181 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 240
Db	187 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 246	RT	187 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 246
QY	241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 300	RL	241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 300
Db	247 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 306	DR	247 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 306
QY	301 TKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 360	DR	301 TKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 360
Db	307 AKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 366	DR	307 AKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 366
QY	361 TRYNDSHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 420	DR	361 TRYNDSHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 420
Db	367 RYNDTHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 426	DR	367 RYNDTHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 426
QY	421 IYKKNATSLRLRRDRNAPNRKITLDDTTAKTSVSVQFAMQLFYDHIQTHINDMFSR 480	DR	421 IYKKNATSLRLRRDRNAPNRKITLDDTTAKTSVSVQFAMQLFYDHIQTHINDMFSR 480
Db	427 IISNNATSVRLRRDLHRASSGKAILNDSVVIIRSTASVQFAMQLFYDHIQAHINEMFSR 486	DR	427 IISNNATSVRLRRDLHRASSGKAILNDSVVIIRSTASVQFAMQLFYDHIQAHINEMFSR 486
QY	481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 540	DR	481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 540
Db	487 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 546	DR	487 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 546
QY	541 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 600	DR	541 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 600
Db	547 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 606	DR	547 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 606
QY	601 GSGYALFENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660	DR	601 GSGYALFENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660
Db	607 GDGYAFYENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 666	DR	607 GDGYAFYENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 666
QY	661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 716	DR	661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 716
Db	667 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 714	DR	667 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 714
QY	717 ELTDLINAKAMTSLASLDQYAK-----TEASLSAYSEAEVTN-----NNLNA--- 758	DR	717 ELTDLINAKAMTSLASLDQYAK-----TEASLSAYSEAEVTN-----NNLNA--- 758
Db	715 VVGAAGAIVSTVSGISAFMSPNPGALAIGLIVIAAGIAFLAYRVVKNLKSNNPKALYP 774	DR	715 VVGAAGAIVSTVSGISAFMSPNPGALAIGLIVIAAGIAFLAYRVVKNLKSNNPKALYP 774
QY	759 -TLEQLMAKTNLESAINQANTDKTTFDNEHNLVE-----AYALKYTLBORAT 807	DR	759 -TLEQLMAKTNLESAINQANTDKTTFDNEHNLVE-----AYALKYTLBORAT 807
Db	775 MTEELK-----DQATRPRTDGS--DSELMASIDERKLEAAREMIKYMALVSAERHOK 826	DR	775 MTEELK-----DQATRPRTDGS--DSELMASIDERKLEAAREMIKYMALVSAERHOK 826
QY	808 NL-----EGLSSTAYNOIRN---NLVDLYNKASSLITKTLDPL 842	DR	808 NL-----EGLSSTAYNOIRN---NLVDLYNKASSLITKTLDPL 842
Db	827 KLKKKRGTAIILSDHLSNRLMNGHRKYDKLNDTSETDDEI 869	DR	827 KLKKKRGTAIILSDHLSNRLMNGHRKYDKLNDTSETDDEI 869
RESULT 10			
Q69408			
ID	Q69408	PRELIMINARY;	PRT; 864 AA.
AC	Q69408;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	GB homolog		
OS	Measgrid herpesvirus 1 (herpesvirus of turkeys).		

Query Match 54.3%; Score 3048.5; DB 12; Length 864;

Best Local Similarity 68.0%; Pred. No. 5.9e-137;

Matches 600; Conservative 90; Mismatches 132; Indels 61; Gaps 11;

QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVTSREVSSVOLSEESTFYLCPPPVGVSTVRL 60

Db 1 MKYFNLSFLFPLPILLSIATSEIKLPNTVAREIVSGIQLSEDETFYVCPVPVGSIVRL 60

QY 61 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 120

Db 61 EPRKCPERKATEWEGEGIAIFLKENISPYKFKVLYLYKNIQTWTGTTTQITNRYT 120

QY 121 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 180

Db 121 DRTPVSEITDLDGKGCSSKARYLRNNVYVAFDRDAGEKOVLLKPSKENTPESRAW 180

QY 181 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 240

Db 181 HTTNETYTVGSPWYRTGTSVNCIVEEMDARSAPPYTFAMANGDIANKSPYGYTPTD 240

QY 241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 300

Db 241 AAAPMGYPQDNFKQLDYSFMDLDRKRRKASLPVRKNFLITSHFTVGVGDWAPKTRVCSM 300

QY 301 TKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 360

Db 301 AKWKEVTEMLRATVNGRYRFRMARELSATFISNTTDPNRIILGQCICKREAAAIQIFR 360

QY 361 TRYNDSHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 420

Db 361 RYNDTHVKGHVQYFALGGFIVAYQPVLSKSLAHMYLRELRDNRDTEMLDLVNKHHA 420

QY 421 IYKKNATSLRLRRDRNAPNRKITLDDTTAKTSVSVQFAMQLFYDHIQTHINDMFSR 480

Db 421 IISNNATSVRLRRDLHRASSGKAILNDSVVIIRSTASVQFAMQLFYDHIQAHINEMFSR 480

QY 481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 540

Db 481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAAVSSCTAIDASVTL 540

QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 600

Db 541 QNSMRVITSTNTCYSRPLVLFYSGENQNGIQQGQNNELLPLEAVEPCSAHRRYFLF 600

QY 601 GSGYALFENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660

Db 601 GDGYAFYENYFVKWVDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660

QY 661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 716

Db 661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKDPANNNGOTQL-BAARM 708

QY 717 ELTDLINAKAMTSLASLDQYAK-----TEASLSAYSEAEVTN-----NNLNA--- 758

Db 709 VVGAAGAIVSTVSGISAFMSPNPGALAIGLIVIAAGIAFLAYRVVKNLKSNNPKALYP 768

216 TSVNCIVEEDARSVPYDSFAISTGDIHMGFFELR-DGAHTEYISYSTORFOQIEY 274
260 FSNLDLRRKASLPKRNFLITSHETVGVWDWAPKTRVCSMTKWKVEVTEMLRATVNGRYR 319
275 YPTDLTRLQGLAPVSRNFLTQHTVTAVNWPVKIREVCTLAKWREIDEIRDEYKGSYR 334
320 FWARELSATFISNTTFFDNRILLGCIKREAEAEIQIFRKYNDSHVVKVGHVQFLAL 379
335 FTAKSISATFISDTTQDFDKLSDCAKREAEIAIDKIYKKYKTKHTQTGELETYLAR 394
380 GGFIVAYQVLSLAHMYLRELMNRDRTDMLDVLNNKHAIVKKNATSLSLRRDIRNA 439
395 GGFIAFRPMISNELAKLYINELVRSNRVVDLSLLNPSVRGAKRRRVEENKSKRNI 454
440 PNRKITLDDTTAIKSTSSVQFAMQLFYDHIQTHINDMFSRIAIDAEVTLQNSMRVITSTNTCYSRPLV 499
455 EGGIENVNNTIKTSSVHFAMQLFYDHIQSHVNMELSRITAWCNLQNKERTLWNEV 514
500 IKINPSATASATLGRVAAKMLGDVAVSSCAIDAESVTLQNSMRVITSTNTCYSRPLV 559
515 MKLNPTSVASVAMDQVSARMGLDVLAVTQCVNIGSSVFIQNSMRVLGTTTCYSRPLI 574
560 LFSYGENOCN-IOGQGENNELLPVLEAVEPCSAHRRYFLGSGYALFENYFKMVA 618
575 SFKALENSTNYIEGQGENNELLVKRLLEPCTANHKRYFKFGVDYVYFENYAVRKVPL 634
619 ADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGLDYAEVARRNOLHELKFDYDINK 678
635 NEIEMISAYDLNITLLEDEILPLSVYTKKEELRDVGLDYAEVARRNOLHELKFDYDINK 694
679 VIEVTNYAGIQEFGCMSTTKKDANPNNGQTOLEAARMELTDLINAKAMTSLASQDYAK- 737
695 VVKVDNNVIMRGIANFQGLDVGAGFGKVLGAA-----NAVIVTSGVSFLNN 746
738 -----IEASLSAYSEAEVNNNLNATLEQLKMAKT-NLESAINQANTDKTTFDN 786
747 PGALAVGLLILAGLFAFLAYRYVSKLSNPKALYPVTRNLSKESVNGSNGNSDGE 806
787 EHPNLVE-----AYKALKTTLEQR-----ATNLEGLS-----STAY 817
807 ENDDNIDEKLOAKEMIKYMSLVSAEQEHAIKKNSGPALLASHITNLSLKHGRPKY 866
818 NQIRN 822
867 KRLKN 871

RESULT 13

Q84735 PRELIMINARY; PRT; 881 AA.
AC Q84735;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein B equivalent.
GN GB.
OS Phocine herpesvirus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=45104;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PB 84;
RC Harder T.C., Osterhaus A.D.M.E.;
RT Identification, sequence analysis and immunogenicity of a marine
RT mammal herpesvirus (phocid herpesvirus-1) glycoprotein B equivalent
RL gene expressed in insect cells.*;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68147; CAA93272.1;
DR InterPro: IPR000234; Glycoprot.B.
DR Pfam: PF00606; Glycoprotein.B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.

SQ SEQUENCE 881 AA; 99574 MW; FCB5E68EB35CC162 CRC64;
Query Match 34.3%; Score 1930; DB 12; Length 881;
Best Local Similarity 45.4%; Pred. No. 8.5e-84;
Matches 382; Conservative 151; Mismatches 238; Indels 70; Gaps 13;
QY 20 NSPSTQ-NVTREVSVSSQL-SEESTFYLCPPVGVGSIVIRLEPPRKPEPRKATEWGE 77
DB 38 NSSSTELNDMDREILGESQSDDTATFMCPPSGSTLVRLEPRACPKNYKLGNFTE 97
QY 78 GIALIFKENISPKFKVLYKNIQIOTWTGTGTQITNRYTDRTPVSEETITDLIDGK 137
DB 98 GLAVIFKENISPKFKANIYKNIITVWSSSYAVTNMHDRIPIKVOEITELIDRR 157
QY 138 GRCSKARYLRNNVYAEFRDAGEKQVLLKPSKFTPESSRAHHTTNETYTVWGSPIWR 197
DB 158 GMCLSKADYIRNNYEFTAFDKDEDPREMLKPSKFTPGSRGHTTNDIYTKIGSPGYR 217
QY 198 TGTSVNCIYBEMDARSVPFYSYPAMANGDIANSPIYGLSPPEAAEPMGYPQDNFKOLD 257
DB 218 TGTSVNCIYBEMDARSVPFYSFGISTGDIHMSPPFGLR-DGAHTEHTSYSDNRFQIE 276
QY 258 SYFSDMLDRKRSKASLPVKNELITSHFTVGDWAPKTRVCSMTKWKVEVTEMLRATVNGR 317
DB 277 GYIPDLDRQLQVGGVSRNFLTQHTVTAVNWPVKIREVCTLAKWREIDEIRDEYKGS 336
QY 318 YRFMARELSATFISNTTFFDNRILLGCIKREAEAEIQIFRKYNDSHVVKVGHVQYEL 377
DB 337 YRFTAKSISATFISDATQFDINRVKLSDCAKREATEAIDKIYKKNYKTKHTQGELETYL 396
QY 378 ALGGFTVAYQVLSLAHMYLRELMNRDRTDMLDVLNNKHAIVKKNATSLSLRRDR 435
DB 397 ARGGFTIAPRPMISNELAKLYINELARSERIVDLNALLNPSHVGGRKRSIETETLGRS 456
QY 436 IRNAPNRKTTLDDTTAIKSTSSVQFAMQLFYDHIQTHINDMFSRIAIDAEVTLQNSMRVITSTNTCY 495
DB 457 KRDVGGVQVNNATILKTTSSIHFAQLFYDHIQSHVNMELSRITAWCNLQNKERTL 516
QY 496 WHEGIRKINPSATASATLGRVAAKMLGDVAVSSCAIDAESVTLQNSMRVITSTNTCY 555
DB 517 WNEVMKLNPTSTITMDQVKSARLLGDVIAVTCVNISSNVFIQNSMRVTGTTTCYS 576
QY 556 RPLVLFSGYENOCN-IOGQGENNELLPVLEAVEPCSAHRRYFLGSGYALFENYFVK 614
DB 577 RPLISFKALENSTDYIEGQGENNELLVDRKLEPCTANHKRYFKFGVDYVYFENYVIR 636
QY 615 MYDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGLDYAEVARRNOLHELKFY 674
DB 637 KVPNLNEIEMISTVVDNLNITLLEDEILPLSVYTKKEELRDVGLDYAEVARRNOLHELKFY 696
QY 675 DINKVIEVDTN-----YAGIQEFGCMSTTKKDANPNNGQTOLEAARMELTDLINA 724
DB 697 DIDSVVKVDNNLIMRGMLTFFQGLDVGAA-----GFGKVLGAA-----NA 738
QY 725 KAMTSLASQDYAK-----IEASLSAYSEAEVNNNLNATLEQL-----KMAKT 768
DB 739 VISTVSGISSFLNPFALAVGLLILAGLFAFLAYRYVSKLSNPKALYPVTRNLSK 798
QY 769 NLESAINQANTDKTTFDNE-----HPNLVEA-----YKALK-----TTLEORATN 808
DB 799 SSKEKIGDGEDGDEDEDEKLSQAKEMIKYMTLISAMEKQEHKAMKNSGPAILANRVAN 858
809 L 809
859 L 859
RESULT 14
ID Q90050 PRELIMINARY; PRT; 948 AA.
AC Q90050;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

Db 745 RRNLQHALKFYDIDSIVRVDDNLVIMRGMAFFQGLGDYVGAGFGKVGVLGAASAVISTVSG 804
 Qy 700 KDANPNNGQQLSEARMLTDLNA-----KAM---TLASLDQYAKTEAS 741
 Db 805 VSSFLNPNFGALVAGLLILAGIVAAFLAYRYSRLRANPMKALYPVTVTRNLKQTAKSPAS 864
 Qy 742 LSSAYSSEATVNNLNATLEOLAKMAK--TNLESAINQANTDKTTFDNEHPNLVEAY---KA 797
 Db 865 TAGGDSDPGVDDFDEKLMQAREMIKMSLVSAWEQ-OEHKAMKKNGPAILTSHLTNMA 923
 Qy 798 LKT--TLEQRTNLEGLSSPAYNOI 820
 Db 924 LRRRGPKYQRLNLDGDDTETNLV 948
 RESULT 15
 Q9QAP5 PRELIMINARY; PRT; 919 AA.
 ID Q9QAP5 AC Q9QAP5
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Glycoprotein B.
 OS Caprine herpesvirus 1 (goat herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae.
 ON NCBI_TaxID=39944;
 RX MEDLINE=99221732; PubMed=10203465;
 RP SEQUENCE FROM N.A.
 RC STRAIN=E/CH;
 RA Ros C., Belak S.;
 RT "Studies of genetic relationships between bovine, caprine, cervine,
 RT and tangierine alphaherpesviruses and improved molecular methods for
 RT virus detection and identification";
 RL J. Clin. Microbiol. 37:1247-1253(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E/CH;
 RA Ros C., Belak S.;
 RT "Characterization of the glycoprotein B gene from ruminant
 RT alphaherpesviruses";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF078728; A046114.2; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD00693; Glycoprot_B; 1.
 SQ SEQUENCE 919 AA; 100172 MW; A7DDCF03D13B4B66 CRC64;
 Query Match 32.1%; Score 1805; DB 12; Length 919;
 Best Local Similarity 51.8%; Pred. No. 7,8e-78;
 Matches 354; Conservative 208; Mismatches 24; Gaps 7;
 Qy 23 PSTQNVTS-----REWSSVOLSEESTFYLCPPVGVSTVIRLEPPRCPEPRK 71
 Db 83 PSADGATAAGNETAADDVRAVLREAAGEHSQFFVCPSPSGATVVR LAPGRCPPEYL 142
 Qy 72 ATEWCEGIALFKENISPYEFKVTLYKNNIIQTWTGTTYRQITNRYTDRTPVSIEIT 131
 Db 143 GRNTEGIGIYKENIAPYFKAHYKNNVIVTTWSGSTYAAATNQHTDRVPVGSVIT 202
 Qy 132 DLIDGKRCSSKARYLNNVYVEAFDRDAGEKQVLLKPSKENTPESRAHWTNTETTYWVG 191
 Db 203 DVVDKWRCLSKAEYLSGRKRVAFDRDADPWEAPLPSRLNAGARGWHITDEVHTVVG 262
 Qy 192 SPWYRTGTSVNCIEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMGYPOD 251
 Db 263 SVGLYRTGTSVNCIEEVEARSVPYDSFALSTGDIYMSPFYGLR-EGAREHTSYSPD 321
 Qy 252 NFKQLDSYFSDLDKRRKASLPVKNRNLITSHFTVGVGDWAPKTYFVCSMTWKKEVTEMLR 311
 Db 322 RFQQIEGYYSKDLTSGRRAGDPVSRNFURTHQVTVAMDVWPKRKNVCSLARWAAEEMLR 381

QY	312	ATVNGRYREMARLSATFISNTTEPDNRIILGOCIKREAEAAIEQIFETKYNDSHVKVG	371
Db	382	DESRGNRYRETARALSATFVSDTHAFSLQHVPLSDCVLEDAEAEVERVYRERYNGTHVLG	441
QY	372	HVOYFLALGGFTVAYQPVLSKSLAHMYLRELMRDNRTDEMLDLVNKKHAIYKKKATSLSR	431
Db	442	KDVTYLARGGFVVAFRPLLSNELAKLYLQELARSNT---LDGVLGPR--LPAGAAPAGA	496
QY	432	LRRDIRNAPNRKITLDDTTAKSTSSVQFAMLOFYDHIQTHINDMFSFIATMCELQNR	491
Db	497	LRARARAAFGGG---GGAGRVTVTVSSAEFAALQFTYDHIQDHVNAMFSRLATSWCILLQNK	553
QY	492	ELVLWHEGJKINPSATASATLGRVRAAKMIGDVAAVSSCTADAESVTLQNSMRVITSTN	551
Db	554	ERALWAEAKLNPSSAAASALNRRAARMIGDAMVTCRELGAGRVFTENSMR-MPGGA	612
QY	552	TCYSRPLVFSYGENQGNIOGQGENNELLPTLEAVEPCSANHRRYFLFGGYALFENYN	611
Db	613	ACYSRPLVSFAYGNESEPLEGGQGEDNELLSSDLVEPCAAHRRYFREGADYVYENYA	672
QY	612	FVKMVDADIQIATFVELNLTLLEDREILPLSVYTKELRDVGVLDDYAEVARRNOLHEL	671
Db	673	YVRRVPLTEIETISTFVDLNLVLEDRFPLEVYTRAEADTGLIDYSEIQRRNOLHEL	732
QY	672	KEYDINKVIEVDNYA---GLQEF	692
Db	733	KEYDIDRVVKLDSNMVIMRGLASF	756

Search completed: October 8, 2003, 17:07:05
Job time : 97.4809 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:35 ; Search time 19.5547 Seconds
(without alignments)
2611.703 Million cell updates/sec

Title: US-09-147-052-4

Perfect score: 5619

Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNENADKIPGRPGTFL 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3601	64.1	865	1	VGLB_HSVMD
2	1785.5	31.8	980	1	VGLB_HSVEA
3	1779.5	31.7	980	1	VGLB_HSVB
4	1775.5	31.6	868	1	VGLB_VZVD
5	1775	31.6	913	1	VGLB_PRVIF
6	1761.5	31.3	979	1	VGLB_HSVL
7	1760.5	31.3	919	1	VGLB_HSVE4
8	1741.5	31.0	980	1	VGLB_HSVE1
9	1736	30.9	904	1	VGLB_HSV2H
10	1735	30.9	904	1	VGLB_HSV23
11	1731.5	30.8	932	1	VGLB_HSVBC
12	1729.5	30.8	920	1	VGLB_HSVSM
13	1714.5	30.5	885	1	VGLB_HSV2S
14	1710	30.4	917	1	VGLB_HSVB2
15	1705.5	30.4	903	1	VGLB_HSV1F
16	1705.5	30.4	904	1	VGLB_HSV11
17	1696.5	30.2	904	1	VGLB_HSV1P
18	1693.5	30.1	904	1	VGLB_HSV1K
19	1675.5	29.8	933	1	VGLB_HSVA1
20	1604	28.5	928	1	VGLB_HSVBP
21	1337.5	23.8	883	1	VGLB_ILTVS
22	1325	23.6	883	1	VGLB_ILTV6
23	1324	23.6	873	1	VGLB_ILTVT
24	778.5	13.9	907	1	VGLB_HCMVT
25	758	13.5	830	1	VGLB_HSV6U
26	756	13.5	906	1	VGLB_HCMVA
27	750	13.3	830	1	VGLB_HSV6
28	741	13.2	822	1	VGLB_HSV7J
29	739	13.2	830	1	VGLB_HSV62
30	738	13.1	901	1	VGLB_GPCMV
31	737	13.1	928	1	VGLB_MCMVS
32	717	12.8	854	1	VGLB_RHCM6
33	709.5	12.6	857	1	VGLB_EBV

34 700.5 12.5 944 1 VGLB_HSVT2
35 695 12.4 808 1 VGLB_HSVSA
36 185 3.3 2660 1 YEEJ_ECO57
37 176 3.1 3712 1 LMA_DROME
38 172 3.1 1085 1 CUT7_SCHPO
39 170 3.0 1251 1 RBP2_PLAVB
40 168.5 3.0 1433 1 CAT8_YEAST
41 164 2.9 1398 1 MHP1_YEAST
42 163 2.9 1093 1 TMF1_HUMAN
43 161 2.9 1630 1 MSP1_PLAFK
44 161 2.9 1639 1 MSP1_PLAFW
45 160.5 2.9 1225 1 Y309_MYCSE

ALIGNMENTS

RESULT 1

VGLB_HSVMD
ID VGLB_HSVMD STANDARD; PRT; 865 AA.
AC P18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293086; PubMed=2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Blinn M.M., Doel T., Milne B.;
RT "Nucleotide sequence and characterization of the Marek's disease virus homologue of glycoprotein B of herpes simplex virus.";
RL J. Gen. Virol. 70:1789-1804(1989).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D13713; BAA02866.1;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; GlycoproteinB; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Signal; Glycoprotein; Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 865 GLYCOPROTEIN B.
FT DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 683 700 POTENTIAL.
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 732 752 POTENTIAL.
FT DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 98091 MW; B30E93C1AC65C6C3 CRC64;

Query Match 64.1%; Score 3601; DB 1; Length 865;

Best Local Similarity 85.9%; Pred. No. 9,1e-175;

Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;

QY 1 MHYFRNCIFFLIVLYGTNSSPQTNTSVRSVSSVOLSEESTFYLCPPVGVSTVRL 60

|||||

1	DB	MHYFRNCIFFLVILLYGTNSPQNTVSRVWSSVQLSEBESTFYLCPPVPVGSVTHKL	60
61	QY	EPKRKCEPRKATEWEGEIALFKENISPYAFKVTLYYKNIIQTWTGTTTYQITNRYT	120
61	DB	EPKRKCEPRKATEWEGEIALFKENISPYAFKVTLYYKNIIQTWTGTTTYQITNRYT	120
121	QY	DRTPVSBEEITDLIDGKRCSSKARYLNNVYVEAFDRDAGBKQVLLKPSKFNTPESRAW	180
121	DB	DRTPVSBEEITDLIDGKRCSSKARYLNNVYVEAFDRDAGBKQVLLKPSKFNTPESRAW	180
181	QY	HTTNETYVWGSPPWYRTGTSVNCIVEMDARSFPYSYFAMANGDIANISFYGLSPPE	240
181	DB	HTTNETYVWGSPPWYRTGTSVNCIVEMDARSFPYSYFAMANGDIANISFYGLSPPE	240
241	QY	AAAPMGYPQDNFKOLDSYFSMDLDRRKASLPVKRNFLLTSHFTVGDWAPKTTTRVCSM	300
241	DB	AAAPMGYPQDNFKOLDSYFSMDLDRRKASLPVKRNFLLTSHFTVGDWAPKTTTRVCSM	300
301	QY	TKWEVTEMLRATVNGRYRFWARELSAFFFISNTTFEDPNRIILGOCIKREAPAAIEQIFR	360
301	DB	TKWEVTEMLRATVNGRYRFWARELSAFFFISNTTFEDPNRIILGOCIKREAPAAIEQIFR	360
361	QY	TKYNDSHVKVGHVQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDEMLDLYNNKHA	420
361	DB	TKYNDSHVKVGHVQYFIALGQFIVAYQPVLSKSLAHMYLRELMRDNRDEMLDLYNNKHA	420
421	QY	IYKKNATYSLSLRDIRNAPNRKTTLDTTAIKSTSSVQFAMQLFYDHIQTHINDMFSR	480
421	DB	IYKKNATYSLSLRDIRNAPNRKTTLDTTAIKSTSSVQFAMQLFYDHIQTHINDMFSR	480
481	QY	IATAWCELONRELVLWHEGINKNPSATASATLGRVAAKMGDVAASVSCCTAIDAESVTL	540
481	DB	IATAWCELONRELVLWHEGINKNPSATASATLGRVAAKMGDVAASVSCCTAIDAESVTL	540
541	QY	QNSMRVITSTNTCYSRPLVLFPSYGENQGNIOQLGENNELLPTLEAVPSCSANHRRYFLF	600
541	DB	QNSMRVITSTNTCYSRPLVLFPSYGENQGNIOQLGENNELLPTLEAVPSCSANHRRYFLF	600
601	QY	GSYGALFENYFVWMDAADQIASTFVELNLTLLDREIPLSVYTKKEELRDVGVLDYA	660
601	DB	GSYGALFENYFVWMDAADQIASTFVELNLTLLDREIPLSVYTKKEELRDVGVLDYA	660
661	QY	EVARRNQLHELKFYDINKVIBVDNTYA--GLQFEGCMSITKKDANPNNGQTOL-EAARM	716
661	DB	EVARRNQLHELKFYDINKVIBVDNTYA--GLQFEGCMSITKKDANPNNGQTOL-EAARM	716
717	QY	ELDTLINAAMWTLASLDQYAK-----IEASLSSAYSEAEVTN-----NNLNA---	758
709	DB	VVWGAAGAIYSTIGSVAFMSNPFCAALIGLIIAGLVAAFLAYRYVNNKLKSNPMKALYP	768
759	QY	TLLEOLKMAKTNLESANQANTDKTTFDNEHPNLVEA-----YKALKTTLSQRATNL	809
769	DB	MTTEVLKAQATRELHGEESDOLSTSIDER--KLEEFAREMIKYMALVASAERHEKKL	823

RESULT 2

RESOLUT_2	VGLB_HSVEA	STANDARD;	PRT;	980 AA.
ID	VGLB_HSVEA			
AC	PI8551;			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor (Glycoprotein 14).			
DE	GB OR GP14 OR 33.			
GN				
OS	Equine herpesvirus type 1 (strain AB1) (EHV-1).			
OS	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OC	NCBI_TaxID=10328;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	Bonasa W.A., Elton D.M., Stocks J.M., Killington R.A.,			
RA	Meredith D.M., Halliburton I.W.			
RL	Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.			

```

Db      620 NPSAIVSATLDERVAARVLGDVIAIHCACAKLEG-NVYLQNSMRSDS-NTCYSRPPVFT 677
QY      563 Y---GENQNTQOGQGENNELLPTEAVEPCSAHRRYFLFGSGYALFENFYKMDAA 619
Db      678 ITKNANRGSEGOGLGEENEFTKRLIEPCALNOKRYFKGKVVYVNTFVRKVPPT 737
QY      620 DIQASTFVELNLTLLEDEILPLSVYTKBELRDVGVLDAEVARNQLHKLKYDINKV 679
Db      738 EIEVISTVELNLTLLEDEILPLSVYTKBELRDVGVLDAEVARNQLHKLKYDINKV 797
QY      680 IEVDTN-----YAGL-----QEFGCMSTIKKDA-----NPNNGQT--- 709
Db      798 VVNDTAVIMOGIASFFKGLGKVGANGVTIVLAAGAVSVSVGIASFLNPPFGLAIGL 857
QY      710 -----OLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAEV 752
Db      858 LVIAGLVAAPFAFYVMQIRSNPKALYPITTKALKNKAKTSYQGNEDDGDGDFDEAKLE 917
QY      753 N-----NNLNATLEOLKKA-KTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803
Db      918 EAREMIKYSMSVALEKQEKKAIRKNSGVGLIASVSKLALRRRGPY-----TRLQ 969
QY      804 QRAT 807
Db      970 QNDT 973

RESULT 3
VGLB_HSVBEB
IDB_HSVBEB STANDARD; PRT; 980 AA.
AC P28922;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR gp14 OR 33
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520, 10330;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=AB4p;
RC MEDLINE=92295566; PubMed=2157895;
RX Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Kentucky D;
RX MEDLINE=90219246; PubMed=2157895;
RA Guo P.X., Goebel S., Perkus M.E., Taylor J., Norton E., Allen G.,
RA Lanquet B., Desmetre P., Paolletti E.;
RT "Coexpression by vaccinia virus recombinants of equine herpesvirus 1
RT glycoproteins gp13 and gp14 results in potentiated immunity.";
RL J. Virol. 64:2399-2406(1990).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; M86664; AAB02468.1; -
DR EMBL; M35145; AAA46067.1; -
DR PIR; G36798; VGBEC6.
DR InterPro; IPR00234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein_B; 1.

```

```

DR      ProDom: PD000693; Glycoprot_B; 1.
KW      Glycoprotein; Transmembrane; Signal.
FT      SIGNAL 1 86
FT      CHAIN 87 980
FT      DOMAIN 87 852
FT      TRANSMEM 853 870
FT      DOMAIN 871 980
FT      CARBOHYD 165 165
FT      CARBOHYD 275 275
FT      CARBOHYD 380 380
FT      CARBOHYD 423 423
FT      CARBOHYD 497 497
FT      CARBOHYD 514 514
FT      CARBOHYD 515 515
FT      CARBOHYD 560 560
FT      CARBOHYD 727 727
FT      CARBOHYD 749 749
SQ      SEQUENCE 980 AA; 109805 MW; A6DDFA8CA5550FF5 CRC64;

Query Match 31.7%; Score 1779.5; DB 1; Length 980;
Best Local Similarity 41.9%; Pred. No. 1.3e-82;
Matches 379; Conservative 132; Mismatches 260; Indels 133; Gaps 17;

QY      19 TNSPSTQNVTS-----REVSSVQLSEESTFYLCPPPVGST 56
Db      88 TTPSPPTSTPTSMTHSGTVDPTLLPTTDPDLRAVRESGILAEADGDFYTCPPPTGST 147
QY      57 VIRLEPPKCPKPRKATEWEGGIALFKENISPYKFKVLYKNIQTTTGTGYROI 116
Db      148 VRIEPPCTPKFDLGRNFTGIAVFKENIAPYKFRANVYKIDIVTRVWKGSHTSLS 207
QY      117 NRYTDRTPVSTIEITDLIDGKRCSSKARYLRNNVYVDFADRDAGEKQVLKPSKFNPE 176
Db      208 DRYNDRVPVSVEEIFGLIDSKGSSKAEYLRDNIIMHAYHDDDEVELDLVPSKFAFG 267
QY      177 SRWHTTNET--YVWGSPIYRTGTSVNCIVEEMDARSVPYFAMANGDIANISFPY 234
Db      268 ARAWOTTNDTTSYGVG-MPWRHYTSTSVNCIVEEARSVPYFSAFSTGDIYASFPY 326
QY      235 GLSPPEAAA--EPMGYPODNFKQLDSYFSDMLDKRRKASLPVKRNLFTSHTFGWDAP 292
Db      327 GL---RAAARLEHNSYAQERFQVGYPRDLDSKLOAEPEVTKNFTTPTVHTVSWNTE 383
QY      293 KTRVCSMTKWKVETMLRATVNGRYRPMARELSATFISNTFEDPNKILIGQCIRDAE 352
Db      384 KKEACTLTWKVEDELVRDEFSGSYRFTIRSISSFTISNTTQFKLESAPLTCVSKAK 443
QY      353 AAIEQIFRTKYNDSHVKVGHVQYFLAGGFTVAYQPVLSKSLAHMYLRELARDNRTDML 412
Db      444 EADISYIAKQYESTHVSFGDVEYIARGGFLIAFRPLMSLNELARLYLNELVRSNTYDLK 503
QY      413 DLVNNKHAIYKKNATISLRRLRRDINRNPRTILD-----DTTA----- 451
Db      504 NLLNPN---ANNNNITRRRSLLSVPEQPTQGVHREQLHLHLKRAVEATAGTSS 559
QY      452 -----IKSTSVQFAMQLQYDHIQTHINDMFSRIATMCELQNLRELVLHHEGKI 502
Db      560 NVTAQLELINTTSIESFAMQLQYDHIQSHVNMELSRITATWCTLQNKERTLNMWVKI 619
QY      503 NPSATASATLGRRAAKMGLDVAASCTADAESVTLQNSMRVTSNTCYSRPLVLS 562
Db      620 NPSAIVSATLDERVAARVLGDVIAIHCACAKLEG-NVYLQNSMRSDS-NTCYSRPPVFT 677
QY      563 Y---GENQNTQOGQGENNELLPTEAVEPCSAHRRYFLFGSGYALFENFYKMDAA 619
Db      678 ITKNANRGSEGOGLGEENEFTKRLIEPCALNOKRYFKGKVVYVNTFVRKVPPT 737
QY      620 DIQASTFVELNLTLLEDEILPLSVYTKBELRDVGVLDAEVARNQLHKLKYDINKV 679
Db      738 EIEVISTVELNLTLLEDEILPLSVYTKBELRDVGVLDAEVARNQLHKLKYDINKV 797
QY      680 IEVDTN-----YAGLQEFG-----CMSITKDKANPNNGQT--- 709

```

```

Db 798 VVNDTAVIMOGIASPFKGLKVGKAVGTVLCAAGAVVSTVSGIASFLNPPGGLAIGL 857
QY 710 -----QLEARMETDLINAKAMTASLODYAKIEASLSAYSAPETV 752
Db 858 LVIAGLVAFFAYRYVMYIRSNPKALYPITTKALKNAKTSYQONEEDGSDFDKALE 917
QY 753 N-----NNLNATLEQLKMA-KTNLESAINQANTDKTFDNEHNLVYAYKALKTTLE 803
Db 918 EAREMKYMSVSALEKQEKKAIKNSGVGLIASVSKLARRGPKY -----TRLQ 969
QY 804 QRAT 807
Db 970 QNDT 973

RESULT 4
VGLB_VZVD
ID VGLB_VZVD STANDARD; PRT; 868 AA.
AC P09257;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein II).
GN 31.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1739-1816(1986).
CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04370; CAA27914.1; -
CC FIR; E27214; VGBE31.
CC InterPro; IPR000234; Glycoprot_B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD000693; Glycoprot_B; 1.
CC Glycoprotein; Transmembrane; Signal.
CC SIGNAL 1 ?
CC ? 868 GLYCOPROTEIN B.
CC DOMAIN ? 679 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 680 695 POTENTIAL.
CC TRANSMEM 701 720 POTENTIAL.
CC TRANSMEM 724 744 POTENTIAL.
CC DOMAIN 745 868 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 440 440 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 557 557 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 868 AA; 98066 MW; 588BE1346F4FF902 CRC64;
Query Match 31.6%; Score 1775.5; DB 1; Length 868;
Best Local Similarity 50.3%; Pred. No. 1.7e-82;
Matches 335; Conservative 122; Mismatches 186; Indels 23; Gaps 7;

QY 31 REVSSVQLSEESTYLCPPVGVSTVIRLEPPKCPKPRKATWEGEGTAILFKNISPY 90
Db 41 REAIHKSQDAETKPTFYVCPPTGTIVRLEPTRTCPDYHLKNETEGIAVYKKNIAAY 100

```

```

QY 91 KEKVTLYKNIQTWTGTYRQITNRYTDRTPYSIEBITDLIDGKRCSSKARYLRNN 150
Db 101 KFAATVYKDVIVSTAWAGSSYQITNRYADRPVIPVSEITDTIDTDFGKSSKATVRRN 160
QY 151 VVFEAFDRDAGEKQVLLKPKSENTPESSRAWHTTNETYTVMGSPWIVRTGTSVNCIYEEMD 210
Db 161 HKVEAFNEKDPQDMPLIASKYNSVCSKAHTTNDTYMVAGTPTGYRTGTSVNCIIIEVE 220
QY 211 ARSVFPYSYFAMANGDIANSFPYGLSPPEAAEPMGYPDQNFKQLDLSYFSDMDLDRKKA 270
Db 221 ARSIFPYDSFGLSTGDIYMSPPFGLR-DGAYREHSNYAMDRFOHEGYRQORDLDFRALL 279
QY 271 SLPVKRNFLITSHTFVGDWAPKTTVCSCMTAKWETMLRATVNGRYRPMARELSATFI 330
Db 280 E-PAARNFLVTPHLTVGNWPKRTEVCSLVKWREVEDYVRDEYAHNFRFTKTLSTTFI 338
QY 331 SNTTEFDPNRIILGOCIKREAAAEIQIFERTKYNSHVKVGHVQVFLALGGFTVAYQPVY 390
Db 339 SETNEFNQIHLSCVKEEARAIINRIYTRYNSHVRTGDIQIYLAGGGFVVPQPLL 398
QY 391 SKSLAHMYLRELMRDNRDTEMLDLVNNKHAIYKNATSLSLRRDIRNAPNRKITLDDTT 450
Db 399 SNSLARLYLQELVRE-----NTNHS-PQKHPTNTRSRSSV-----PVELRANR 441
QY 451 AIKSTSSYQFAMLOFLYDHIOHINDMFSAIATACELONRELVLWHGKIKINPSATASA 510
Db 442 TTTTSSVEFAMLOFTYHQIEHVNEMLARISSSQCOLQNRERALLSGFLPINPSALAT 501
QY 511 TLGRVAAKMLGDVAASVSCSTDAES-VTLQNSMRVITSTNTCSRPLVLFVSYGENOCN 569
Db 502 ILDRVKARILGDVLSVSNCPGLSDTRILLQNSHRVSGSTTCYSRPLISVLSNGSGT 561
QY 570 IOQGLGENNELPTLEAVEPCSANHRRYELFGSGYALFENFYKVMVDAADIOIASTFVE 629
Db 562 VEGQLGTDELIMSRLDLEPCVANHKRYELFGHHVYVYEDYRYVREIAVHDVGMISTYVD 621
QY 630 LNTLLEDEIILPLSYTKEELRDVGLDYAEVARNQHLKFFYDINKVIEVDNYA-- 687
Db 622 LNTLTLKREFPLQVYITRDELDTGLDYSIQRRNQMSLRFYDIDKVKVQYDSGATM 681
QY 688 -GLOEF 692
Db 682 QGMAQF 687

RESULT 5
VGLB_PRVIF
ID VGLB_PRVIF STANDARD; PRT; 913 AA.
AC P08355;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein gII precursor.
OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87284141; PubMed=3039163;
RA Robbins A.K., Dorney D.J., Wathen M.W., Whealy M.E., Gold C.,
RA Watson R.J., Holland L.E., Weed S.D., Levine M., Gioroso J.C.,
RA Enquist L.W.;
RT "The pseudorabies virus gII gene is closely related to the gB
RL J. Virol. 61:2691-2701(1987).
RP SEQUENCE OF 847-913 FROM N.A.
RX MEDLINE=89279298; PubMed=2543777;
RA Simon A., Mettenleiter T.C., Rziha H.J.;
RT "Pseudorabies virus displays variable numbers of a repeat unit
RT adjacent to the 3' end of the glycoprotein gII gene.";
RL J. Gen. Virol. 70:1239-1246(1989).

```



```

FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 979 AA; 110331 MW; 9A19868B791C5B36 CRC64;

Query Match
Best Local Similarity 31.3%; Score 1761.5; DB 1; Length 979;
Matches 347; Conservative 115; Mismatches 193; Indels 69; Gaps 11;

Qy 19 TNSPSTQNTS-----REVSSVQLSEESTFVLCPPVGST 56
Db 88 TTPSPPTSTPTSMTHSGTVDPTLLPTTDPDLRLAVRESGILAEDEGDFYTCPPPTGST 147

Qy 57 VIRLEPRKCPKPRKATGEGIAILFKENISPYKFKVLYKNIOTTTGTYRQIT 116
Db 148 VRIEPTPCPKFDLGRNFTGIAVFKENIAPYKFRANVYKDIQIVRWKGGSHWLS 207

Qy 117 NRYDTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPE 176
Db 208 DRYNDVPVSEEEIFGLIDSKGCSKAEYLRDMIMHAYHDDDEVELDLVPSKFATP 267

Qy 177 SRAMHTTNET--YTVMGSPWIYRTGTSVNCIVEEDARSVPYSEFAMANGDIANISPFY 234
Db 268 ARAWQNTDNTTSYGVG-MPWRHYTSTVNCIVEEARSVPYDSFALSTGDIYVASFY 326

Qy 235 GLSPPEAAA--EPMGYPDQNKQDLSYFSMDLDRKRRKASLPVKRNFLLTSHFTVGDWAP 292
Db 327 GL---RAAARLEHNSYAQERFQVEGYRPRDLDSKLQAEPEVTXNFITTPHYVSWNTE 383

Qy 293 KTRVCSMTWKVEYTEMRLATVNGRYRPMARELSATFISNTEPDNRILIGQCIKRAE 352
Db 384 KKEACTLTKWKEVDLVRDEFRGSRFTIRSISSYFTSNTQFKLESAPLTCVSKREK 443

Qy 353 AAIEQIPRTKYNDSHVKGHVQYFLALGGFTVAVQPVLSKSLAHMYLRELMDRNTDEML 412
Db 444 EADISYIKQYESTHVSFGDVEYVLARGGFLIAFRPLSNELARLYLNLVRSNRTYDLK 503

Qy 413 DLVANKHAIYKKNATSLSLRRDLIRNAPNRKITLD-----DTTA----- 451
Db 504 NLLNPN---ANNNNNTTRRRSLLSVPEQPTQGVHREQILHRLHRAVEATAGTSS 559

Qy 452 -----IKSTSSVQFAMQLFYDHIQTHINDMFSTRATACELQNRRLVLMHEGIKI 502
Db 560 NVTAKQLELITTSIEFAMQLQFAYDHIQSHVNEMLSRATATACPLQNKERPLNEMVKI 619

Qy 503 NPSATASATLGRVAAKMLGVAAVSSCTADAESVTILQNSMRVITSTNTCYSRPLVFS 562
Db 620 TPSAIVSATLDERVAARVLGVDAIATHCATKTEG-NVYLQNSMRSDS-NTCYSRPPVFT 677

Qy 563 Y---GENQGNTOGQGNELLPTLEAVEPCSAHRRYFFLFGSGYALFENTYFVKMVDAA 619
Db 678 ITKANNRGSIEGOLGEENEIFTERKLEPCALQKRYKFKGVYVYENTYFVKVPPT 737

Qy 620 DIQTASTFVNLNLLDERELPLSVYTKBELRGVLDYAEVARRNQLHEKLYDINKV 679
Db 738 EIEVISTVVELNLLDERELPLSVYTKBELRGVLDYAEVARRNQLHEKLYDINKV 797

Qy 680 IEVD 683
Db 798 VNVD 801

```

RESULT 7

```

VGLB_HSVE4
ID VGLB_HSVE4 STANDARD; PRT; 919 AA.
AC P17472;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Glycoprotein B precursor.
GN GB.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
OS type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID-10333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89125704; PubMed-2915378;
RA Riggio M.P., Cullinane A.A., Onions D.E.;
RT Identification and nucleotide sequence of the glycoprotein GB gene
of equine herpesvirus 4.
RL J. Virol. 63:11123-1133(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26171; AAA46106.1; ALT_INIT.
DR PIR; A31880; VGBEQH.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
DR KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 919 GLYCOPROTEIN B.
FT DOMAIN 41 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 809 POTENTIAL.
FT DOMAIN 810 919 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 566 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 688 688 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 919 AA; 103710 MW; C77E48F26C37BC7B CRC64;

Query Match
Best Local Similarity 31.3%; Score 1760.5; DB 1; Length 919;
Matches 361; Conservative 134; Mismatches 246; Indels 83; Gaps 14;

Qy 42 EESTFYLCPPVPVSTVIRLEPPKCPKPRKATGEGIAILFKENISPYKFKVLYKNI 101
Db 74 EDGDFYTCPPPTGTVVRIEPPKCPKPRKATGEGIAILFKENISPYKFRANVYKDI 133

Qy 102 IQTTTGTGTYRQITNRYDTPVSIIEITDLIDGKRCSSKARYLRNNVYVEAFDRDAG 161
Db 134 VTVKWKGYSTSDRYNDVNPVSVSEEIFTLIDGKRCSSKAEYLRDMIMHAYHDD 193

Qy 162 EKOVLLKPSFNTPESSAWHTTNET--YTVMGSPWIYRTGTSVNCIVEEDARSVPY 219
Db 194 EVELDLVPSKFATPGARAWQNTDNTTSYGVG-MPWRHYTSTVNCIVEEARSVPY 252

Qy 220 FAMANGDIANISPYGLSPPEAAEPGYPQDNFKQDLSYFSMDLDRKRRKASLPVKRNF 279
Db 253 FALSTGDIYVTSYFYLGR-SAAQLEHNSYAQERFQVEGYRPRDLDSKLQAGEPV 311

Qy 280 ITSHTVGDWAPKTTIRVCSMTWKVEYTEMRLATVNGRYRPMARELSATFISNTE 339
Db 312 TTPHYVSWNTEKRIEACTLTKWKEVDLVRDEFRGSRFTIRSISSYFTSNTQFK 371

Qy 340 RIILGQCIKRAEAAIEQIPRTKYNDSHVKGHVQYFLALGGFTVAVQPVLSKSLAH 399

```


Db 372 DAPLTCVSKAKAIDSIRKQYESTHYVFSGDVEYFARGGLIAFRPMISNELARLYL 431
 QY 400 RELMDNRDTEMDLVLNNKHAIYKKNATS-----LSRLRDRIN 438
 Db 432 NELVSRNRTYDLKNLL-NPNANHTNTRRSLLSIPEPTQESLHREILHRLKRAVE 490
 QY 439 APN----RKITLDDTAISTSSVQFAMLOFLYDHIOTINDFESRIATAWCELQNLREL 494
 Db 491 AANSTNSNVTAQLELIKTTSIEFAMLOFAYDHIQSHVNEMLSRATIAWCTLQNKERT 550
 QY 495 LMHEGKINPSATLGRRAAKMLGDVAVSSCTAIDAESVTILQNSMRVITSTNTCY 554
 Db 551 LNNENKVNPSAIVSATLDERVAARVLGDVIAITHCVKIEG-NVYLQNSMR-SSDSNTCY 608
 QY 555 SRPLVLSYGEN---QGNTQOGLGENNELLPLEAVEPCSAHRRYFLFGSGYALFENYN 611
 Db 609 SRPPVFTTITKNANSGRTTEGOLGENEYVTERKLTIEPCAINOKRYFKFGKVVYENYT 668
 QY 612 FYKMDAADIQIATSTFVELNLTLLEDREITLPLSVYTKELRDVGVLDAEVARRNOLHEL 671
 Db 669 YVRKVPPTIEIVISTYVELNLTLLEDREFLPLLEVYTRAELEDGLDYSEIQRRNOLHAL 728
 QY 672 KYVDINKVLEVD-----TNYAGLQEFQCMSTTKKDPANPNQOTOLEAARMELTDL 721
 Db 729 REYDIDSVVNDTAVIMOGIATFFKGLGKVG-----EAVGLIVLGA 770
 QY 722 INAKAWTLASLDYAK-----IPASLSAYSEAEV-----NNLNATLEQLKMA 766
 Db 771 AGAVSTVSGIASFINPNPFGGLAIGLLVIAGLVAFFAYRYVQNLRSNPMKALYPIPTRS 830
 QY 767 KTNLSAIVNAQNTDKTFNEHPNLVEA-----YKALKTTLQOR 805
 Db 831 LANKAKAYGQNDDDTSDTDFEAKLEAREMIKMYMSVALEKQ 874

RESULT 8

VGLB_HSV1 STANDARD; PRT; 980 AA.
 AC P18050;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR G14 OR 33.
 OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10327;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89279217; PubMed=2543744;
 RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
 RA Woodworth L.M.;
 RA "Identification and nucleotide sequence of a gene in equine
 RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
 RT major envelope glycoprotein gB";
 RL J. Gen. Virol. 70:383-394 (1989).
 CC -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00401; BAA00304.1; ALT_SEQ.
 DR PIR; A31241; VGBE2H.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 86
 FT CHAIN 87 980
 FT DOMAIN 87 852
 FT TRANSMEM 853 870
 FT DOMAIN 871 980
 FT CARBOHYD 165 165
 FT CARBOHYD 275 275
 FT CARBOHYD 380 380
 FT CARBOHYD 423 423
 FT CARBOHYD 497 497
 FT CARBOHYD 514 514
 FT CARBOHYD 515 515
 FT CARBOHYD 560 560
 FT CARBOHYD 727 727
 FT CARBOHYD 749 749
 SQ SEQUENCE 980 AA; 109931 MW; 94B69AF82FB6AC2 CRC64;
 Query Match 31.0%; Score 1741.5; DB 1; Length 980;
 Best Local Similarity 41.4%; Pred. No. 1e-80;
 Matches 375; Conservative 133; Mismatches 262; Indels 135; Gaps 19;
 QY 19 TNSSPSTQWTS-----REVSSVOLSEESTFYLCPPPVGST 56
 Db 88 TTPSPPTSTPSTSHSGTVDPTLPTETDPLRLAVRESGLAEDGDFYTCPTPGST 147
 QY 57 VIRLEPPKCPKPRKATENGEGIALFKENISPYKFKVLYYKNIITQTTTGTGTYRQIT 116
 Db 148 VVRIEPPRTCPKFLGRNFTGIAVIFKENIAPYKFRANVYKDVIVTRVWKGYSHTSLS 207
 QY 117 NRYDTRTPYSIEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQV-LKAPSFNT 175
 Db 208 DRYNDRVPVSVBEIFGLIDSKGCKSAEYLRDNIHAYHDEDEVEDELDCRPS-LQLR 266
 QY 176 ESRAWHTTNET--YTVWGSFPIYRTGTSVNCIVEEMDARSVPFYSYFANANGDIANISPF 233
 Db 267 GARAWQTNDTTSYGVW-MPWRYHTSTVNCIVEEVEARSVPYDPSFALSTGDIVTASFP 335
 QY 234 YGLSPPEAAA--EPMGYPDQNFQKLDYSFMDLDRKKASLPVKRNFILTSHTFTVGWDA 291
 Db 326 YGL--RAAAIEHNSYAOERFQVGRPRDLDSKLOAEPEVTKNFITTPHVTVSNNWT 382
 QY 292 PKTTRVCSMTKWKVEMLRATVNGRYRMARELSATFISNTEFDPNRIILGOCIKREA 351
 Db 383 EKKVEACTLTKWKEVDLDEFRGYSIRFIRISSTFISNTTQFKLESAPLTECVSKEA 442
 QY 352 EAAIEQIFRTKYNDSHVKVGHVQYFALAGGFIVAYQFVLSSKSLAHMYLRELMDNRDEM 411
 Db 443 KEAIDSIIKKQYESTHVSDDVEYLLARGFLIAFRPMLSNELARLYNELVRSNRYDL 502
 QY 412 LDLVNKHAIYKKNATSLRLRDRIRNAPNRKITLD-----DTTA---- 451
 Db 503 KNLLNPN-----ANNNTTTRRRSLLSVPEPQTDGVHREQLHRLHRAVEATAGTDS 558
 QY 452 -----IKSTSSVQFAMLOFLYDHIOTINDFESRIATAWCELQNLRELWHEGK 501
 Db 559 SNVTAKQLELIKTTSIEFAMLOFAYDHIQSHVNEMLSRATIAWCTLQNKERTLWDMVK 618
 QY 502 INPSATASATLGRRAAKMLGDVAVSSCTAIDAESVTILQNSMRVITSTNTCYSRPLVLF 561
 Db 619 INPSAIVSATLDERVAARVLGDVIAITHCAKIEG-NVYLQNSMRSDS-NTCYSRPPVTF 676
 QY 562 SY---GENOGNTOGQIGENNELLPLEAVEPCSAHRRYFLFGSGYALFENYFVKWDA 618
 Db 677 TITKANNRGSIQGLGENEYVTERKLTIEPCALNOKRYFKFGKVVYENYTVFRKVPV 736
 QY 619 ADIQTASTFVELNLTLLEDREITLPLSVYTKELRDVGVLDAEVARRNOLHELKFFDYDK 678
 Db 737 TELEVISTYVELNLTLLEDREFLPLLEVYTRAELEDGLDYSEIQRRNOLHALRFDYIDS 796
 QY 679 VIEVDTN-----YAGIQBFG-----CMSTTKKDANPNQOT-- 709
 Db 797 VVNDVTAVIMOGIASFFKGLGKVGAVGLVILGAAGAVSVYSGIASFLNPNPFGGLAIG 856

Qy 710 -----QLEAARMELTDLINAKAMTASLODYAKIEASLSAYSRAET 751
 Db 857 LLVIALGLVAAFFAYRYVMQIRSNPKALYPTTTKALKNAKTSYQNEEDDGSDFDEAKL 916
 Qy 752 VN-----NNLNATLEQLKMA-KTNLESAINQANTDKTTFDNEHNLNVEAYKALKTTIL 802
 Db 917 EAREMIKYMWSVALEKQEKKAIKKNSGVGLTASNVSKLALRRRGPKY-----TRL 968
 Qy 803 EORAT 807
 Db 969 QQNDT 973

RESULT 9
 VGLB_HSV2H STANDARD; PRT; 904 AA.

AC P08666; P89450;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL27.
 OS Herpes simplex virus (type 2 / strain HG52).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 RN NCBI_TaxID=10315;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87071654; PubMed=3024391;
 RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
 RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
 comparison with the corresponding gene of HSV-1.";
 RL Virology 155:322-333(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 2: GH, GG, GI, AND GE.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M14923; AAA66440.1; .
 DR EMBL; Z86099; CAB06752.1; .
 DR InterPro: IPR000234; Glycoprot.B.
 DR Pfam: PF00606; Glycoprotein.B; 1.
 DR Prodom; PD000693; Glycoprot.B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 904
 FT DOMAIN 23 727
 FT TRANSMEM 728 743
 FT TRANSMEM 749 768
 FT TRANSMEM 772 792
 FT DOMAIN 793 904
 FT CARBOHYD 82 82
 FT CARBOHYD 136 136
 FT CARBOHYD 393 393
 FT CARBOHYD 425 425
 FT CARBOHYD 486 486
 FT CARBOHYD 671 671
 FT CONFLICT 92 92
 FT CONFLICT 198 198

FT CONFLICT 308 308 S -> T (IN REF. 1).
 FT CONFLICT 438 438 L -> Q (IN REF. 1).
 FT CONFLICT 568 568 S -> A (IN REF. 1).
 FT CONFLICT 619 620 EL -> DV (IN REF. 1).
 FT CONFLICT 636 636 R -> G (IN REF. 1).
 SQ SEQUENCE 904 AA; 100217 MW; AB050A3AFB4F1066 CRC64;

Query Match 30.9%; Score 1736; DB 1; Length 904;
 Best Local Similarity 43.0%; Pred. No. 1.8e-80;
 Matches 363; Conservative 142; Mismatches 269; Indels 70; Gaps 16;

Qy 20 NSSPSTQNTSVREWSVQVLSSEESTFYLCPPVGVSTVIRLEPPRCPCPEPKKATEWGEI 79
 Db 82 NATVAAGHATLRAHLREIKVENADAQFYVCPPTGATVYVQFEQPRRCPTPEQONTYGI 141
 Qy 80 AILFKENISPYKPKVTLTKYKNIITTTTGTGTQITNRYTDRTPVTSIEITDIDGKR 139
 Db 142 AVYFKENIAPYKFKATMYKDVTVSQWFGHRSQPMGFEDRAPYFEEVIDIKYTKGV 201
 Qy 140 CSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFTPESSRAWHTTNTYTVWGSFWYRTG 199
 Db 202 CRSTAKYVRNNMETAFHRDDHETDMELKPAKATRTSRGWHTTDLKYNPSRVEAFHYG 261
 Qy 200 TSVNCIVEENDASVPYSYFAMANGDIANISFYGLSPPEAAEPMGYPQDNFKOLDYSY 259
 Db 262 TTVNCIVEEDARSVPYDFEVLATGDFVYMSFYGR-EGSHTHTSYAADRKFQVDGF 320
 Qy 260 FMSDLDRKRRKASLPVKNRFLITSHFTVGDWDAKTTTRVCSMTKWKVEYTEMRLATVNGRYR 319
 Db 321 YARDLTTKARATSPPTTNRLLTTPKFTVADWPKPAVCTMTKQWEDENLRAEYGGSPR 380
 Qy 320 FMARELSATFISNTTEDPNRILGQCIKREAAIEQIFRTKYNDSHVKVGHVQFLAL 379
 Db 381 FSDAISITFTTNLTYSLSRVDLDCIGRDAREIDRMFAKYNATHIKVGPQYYLAT 440
 Qy 380 GGFTVAYQPVLSKSLAHMYLRELNRDNTDEMLEDLVNKHAIYKKNATSLSLRRDRNA 439
 Db 441 GGFLIAQPLLSNTLAELYVREYREQ-----DRK-----PRNATP-----APUREA 482
 Qy 440 PNKRITLDDTTATKSTSVQFAMLOFLYDHIQTHINDMFRIATAMCELQNRRLVLMHEG 499
 Db 483 PSANASVE---RIKTTSSIEFARLOFTYNIHQHVNDMLGRIVAMCELQNHLETLWNEA 539
 Qy 500 IKINPSATASNTLGRVAAKMGDVAASVCTAIDAESVTLONSMRVITSTNTCYSRPLV 559
 Db 540 RKLNPNAIASATVGRVRSARMLGDVMAVSTCVVPAPDNVIVQNSMRVSSRPGTCYSRPLV 599
 Qy 560 LFSYGENQGN-IOQLGNNELLPTLEAVPECSANHRRYFLFGSGYALFENYFVKMVA 618
 Db 600 SFYR-EDQGPLIEQLGNNELRLTRDALEPCTVGHRRYFIFGGYVYFEEYAYSHQLSR 658
 Qy 619 ADIQIASTFVELNLTLEDEILPLSVYTKELRDVGVOYAEVARRNQLHELKIFYDINK 678
 Db 659 ADVTVTSTFDLNTITMEDEHFVPLEVYTRHEIKDGLDYTEVQRNQLDLAFADIDT 718
 Qy 679 VIEVDTN---YAGLOEF-----GCMSTTKKDANPNNGNQGLEARMELTDLI 722
 Db 719 VIRADANAAMPAGLCAAFEGMGDLGRVGVGVGVGVSAVSGVSF-----MGNPF 772
 Qy 723 NAKAMTLASLDYAKIEASLSAYSRAETV-----NNLN-----TLBQLKMAKTNLES 774
 Db 773 GALAVGLLVL-----AGLVAFFAFRYVLQLRNPKALYPLTTRELKTSDFPGVGG 825
 Qy 775 NQANTDKTTFDNEHNPVLEA-----YKALKTTLE--QRAYNLEGLSTAYNIQNLVDL 827
 Db 826 GEEGAEGGGFD--EAKLAAREMIRYALVSAMERTEHAKKKGTSALLSSKVTNMVLRK 883
 Qy 828 YNKA 831
 Db 884 RNKA 887

RESULT 10

Db 262 TTVNCVIEEVDARSVPYDEFLVATGDPVYASPPFYGR - EGSHT EHTSYAADRFPQVDGF 320
 QY 260 FSDMLDKRRKASLPVKRNFLLTSFTVCGWDHAPKTRVCSMTKKWEVTEMLRATVNGRYR 319
 Db 321 YARDLTTTKARATSTPTNNLLTPKFTVAMDVAPKRPACTVTKWQEDVEMLAERYGGSPR 380
 QY 320 FMARELSATFSINTTEFPNRIILGQCILKRAEAAIEQIFRTKYNDSHVKVGHVOYFFLAL 379
 Db 381 FSSDAISTTTNLQYSLSRVDLGDCLGRDAREIDRFARKYNATHIKVQPOYYLAT 440
 QY 380 GGFTVAYQPVLSKSLAHMYLRELMRDNRTDEMLDVLNNKHAYIKKNKATSLSLRLRDIRNA 439
 Db 441 GGFLIAYQPLLSNTLAEIYVREYREQ-----DRK-----PRNATP-----APLRA 482
 QY 440 PNRKITLDDTTAISTSSVQAFQMLFYDHIOTHIINDMFSTRATACELQNRRELYVHREG 499
 Db 483 PSANASVE---RIKTTISIEFARLQFTYNHQTVNDMLGRITAVAWCELQNHLETLWNEA 539
 QY 500 IKINPSATASATLGRYAAKMLGDVAASVSSCTAIDAESVTILQNSMRVITSTNTCYSRPLV 559
 Db 540 RKLNPNAIASATVGRYRSARMLGDVMAVSTCVVPADNVIVQNSMRVSSRPCTCYSRPLV 599
 QY 560 LFSYGENQGN-IOGOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENFYKMWDA 618
 Db 600 SFYR-EDQGPLIEGOLGENNELRLTRDALEPCTGHRRYFTFGGCVYVFEYAYSHQLSR 658
 QY 619 ADIOIASTFVELNLTLEDRLPLSVYTKELRDVGLDYAEAVARRNQLHKLKYDINK 678
 Db 659 ADVTVSTFIDLNTIMLEDHEFVLEVPYTRHIEKDSGLDYTEVQRNQLHDLRFADIDT 718
 QY 679 VIEYDTN---YAGIQEF-----GCMSTTKKDDANNNGOTOLEAARMELTDLI 722
 Db 719 VIRADANAMFAGLCATFEGGDLGRAVGKVMGVGVSVASGVSSF-----MSNPF 772
 QY 723 NAKAMTILASQDYAKIBASLSAYSEATV-----NNLNLA-----TLEOLKMAKTNLESAL 774
 Db 773 GALAVGGLVL-----AGLVAAPFAFVYVLQLRNPMKALYPLTTKELKTSDDPGVGGE 825
 QY 775 NQANTDKTTEDNEHPNLVEA-----YKALKTLE--QRATNLEGLSSATYAGINRNVLVDL 827
 Db 826 GEGAEAGGGFD--EAKLAAREMIRYMALVSAMERTEHKARKKGTSSALLSSKVYNNVLRRK 883
 QY 828 YNKA 831
 Db 884 RNKA 887
 RESULT 11
 VGLB_HSVBC
 ID VGLB_HSVBC STANDARD; PRT; 932 AA.
 AC P12640;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein 1 precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
 DE (Glycoprotein 16) (Glycoprotein G130) (Glycoprotein B).
 GN GI OR UL27
 OS Bovine herpesvirus type 1 (strain Cooper).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 ON NCBI_TaxID=10323;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86300884; PubMed=2841484;
 RX Whitbeck J.C., Bello L.J., Lawrence W.C.;
 RT "Comparison of the bovine herpesvirus 1 gI gene and the herpes
 simplex virus type 1 gB gene";
 RL J. Virol. 62:3319-3327(1988).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M21474; AAA46055.1; -
 EMBL; 278205; CAB01598.1; -
 EMBL; AJ004801; CAA06106.1; -
 PIR; A20877; VGBEC.
 InterPro; IPR000234; Glycoprot_B.
 Pfam; PF00606; Glycoprotein_B; 1.
 ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 67
 FT CHAIN 68 932 GLYCOPROTEIN I.
 FT DOMAIN 68 758 EXTRACELLULAR.
 FT TRANSMEM 759 827 POTENTIAL.
 FT DOMAIN 828 932 CYTOPLASMIC.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SEQUENCE 932 AA; 101195 MW; 9DCEAA85C5FC3DA3 CRC64;

Query Match 30.8%; Score 1731.5; DB 1; Length 932;
 Best Local Similarity 49.8%; Pred. No. 3.1e-80;

Matches 341; Conservative 105; Mismatches 218; Indels 21; Gaps 7;

18 GTNSPSQNTSREWSSVLSSEESTFYLCPPVGVSTVIRLEPPKCPPEPRKATEWGE 77
 97 GDRAASPDNSDVAALRLAQAAGENSFFVCPSPGATVVRAPARPCPEIGLGRNTE 156
 78 GIALFRENISPYKFKVLYKNIQTTTGTTRQINRYTDRTPVSEIEITDLIDGK 137
 157 GIGVIYKNIAPYTFKAYI-YKNVITVTWAGSTVAAITNQYDRVPVGMGEITDLVDK 215
 138 GRCSSKARYLNRYVVEAFDRDAGEQVLLKPSFNTPEPESAWITTTETTYWGSPIYR 197
 216 WRCLSKAEYLSGRKRVAFDRDDDDPWEAPLKPAISAPGVKRWHTTDDVYTAGSAGLYR 275
 198 TGTSVNCIVEEMARSVPFYSYFAMANGDIANISPFYGLSPPEAAAEPMGVPODNFKOLD 257
 276 TGTSVNCIVEEARSVPYDSFALSTGDIYNSPFYGLR-EGAHREHTSYSPRFQOIE 334
 258 SYFSMDLDRKASLPVKNRFLTSHFTVGVNDWAPKTRVCSMTKWKKEVTEMLRATVNGR 317
 335 GYKRDMDATGRRLKEPVSRNFRQHTVAVMDWVPKRNKVCSLAKREADEMLRDSRG 394
 318 YRFWRELSATFISNTTEEDPNRILGQCCKREAAAEIQIFRTKYNDSHVKGHVQYFL 377
 395 FRFTARLSATFVSDSTFTALQNPLSDCVIEEAAAEVRYERYNGHTVLSGSLEYL 454
 378 ALGFGIYAYQVLSKSLAHMLRLMDNRDNTDEMLDLVNNKHAIYKKNATSLSLRLDIR 437
 455 ARGGEVFAVRPMLSNELAKLYQLALRSNGTLEGLFAA----AAPKPGP---RRARRAP 507
 438 NAPNRKTYL-----DDTTAISTSSVQFAMQLQFLVDHIOTINDMFSRIATWCBELQ 490
 508 SAPGGPGAANGPAGDGDAGGRVTTVSSAEPAALQFTYDHIQDHVNTMFSRLATSWCLLQ 567
 491 RELVLWHEGKINPSATASATLGRVRAAKMLGDVAVSSCTAIDAESVTLQNSRVITST 550
 568 KERALWAEAKLPASAASALDRRAARMMLGDAMAVTYCHELGEGRVFTIENSMR--APG 625
 551 NTCYSRLVLYSYGNGQNTQGLGNNELLPLEAVEPCSANRRYFLFGSGYALFENY 610
 626 GVCYSRPPVSAFNGNESEPPVEGQLGEDNELLPGRELVEPCETANKKRYFRFGADYVYENY 685
 611 NFVKWDAADIIQIATSTFVELNLTLLDERELPLPLSVYTKBELRDVGLDYAEVARRNQLHE 670

Db 686 AVRRVPLAEVISTFTVDLNLTVLEDFLEFVYTRAEALADTGILDYSEIQRNQLHE 745
 QY 671 LKFYDINKVIEVDNYA---GLOEF 692
 Db 746 LRFYDIDRVKTDGNNAMRGLANF 770

RESULT 12

VGLB_HSVSM
 ID VGLB_HSVSM STANDARD; PRT; 920 AA.
 AC Q04464;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN UL27.
 OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset herpesvirus).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae;
 OC NCBI_TaxID=10353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93228440; PubMed=8385913;
 RA Eberle R., Black D.;
 RT "Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine monkey alpha-herpesviruses.";
 RL Arch. Virol. 129:167-182(1993).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE REQUIRED FOR VIRAL GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M95786; AAA43841.1; -
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 920 GLYCOPROTEIN B.
 FT DOMAIN 29 739 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 740 755 POTENTIAL.
 FT TRANSMEM 761 781 POTENTIAL.
 FT TRANSMEM 784 804 POTENTIAL.
 FT DOMAIN 805 920 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;

Query Match 30.8%; Score 1729.5; DB 1; Length 920;
 Best Local Similarity 49.9%; Pred. No. 3.9e-80;

Matches 332; Conservative 95; Mismatches 209; Indels 29; Gaps 4;

QY 31 REVSVSVQSEESTYLCPPVGVSTVIRLEPPKCPPEPRKATEWGEIATLKFENISPY 90
 Db 109 RESVRIARENATSMFYVCPPTGATVQVFEPRPCPDVAAGKNFTGIAVIFENIAPY 168
 QY 91 KFKVLYYKNIQTTTGTTRQINRYTDRTPVSEIEITDLIDGKRCSSKARYLRNN 150

```

169 KTATMYKEITVQTQSGRIQLTGLYNDRAPVPPFEITDVINAKGLCRSDVTVRSQ 228
151 VVVEAPDRAGKQVLLKPSKNTPSRAWHNTTNYTVWGSFWIYRTQTSVNCIIVEMD 210
229 RRVATYDRDEGWREKVLKPSKTSNRSRGWYTTDRMYAPNAHAGFYKAGTIVNCIIVEE 288
211 ARSVFYSFAMANGDIANSFYGLSPPEAAEPGYQDNFKQLDSYFSDLDKRRKA 270
289 ARSAPYSNFVLTGDFVTVSPFYGLG-EDAHREYNAYSADRPQVDGFFRDLDGSETA 347
271 SLFVRNFIITSHFTVGDWAPKTRVCSMTKWKVEITMLRATVNGRYRFMAELSATFI 330
348 PEVVNRLITTPKFIGWDMKPKDPSVCSVTWKQVEEMRAEYGSTFRFTSSLSATFT 407
331 SNTTEPDNRIILGQCIKREAAEAIQIFRTKYNDSHVGVHVVQVFLALGGFTVAYQPV 390
408 TNVQTPPQRIELSDCVAREAAQADAIYARRYNASHVKGVLQYVLAQGGFLVYQPLI 467
391 SKSLAHMYLRELNRDNTDEMLDLVNKHAIYKKNATSLSLRRDIRNAPNRKIILDDTT 450
468 SNSLAEMYLRE-----AARALEPAPLEPTTAPAE---AAGSRG 502
451 AIKSTSSVQFAMLOFLYDHIQTHINDMFESRIATWCELONRELVLWHEGKINPSATA 510
503 TLTSTQSVFARLQFTYDHIQHVNEMLGRIAAWCOLQNELVLMNEARKLNPAIASA 562
511 TIGRRVRAKMGDVAASCTAIDAESVTLQNSMRVITNTCYSLPLVFSYGENQGI 570
563 TVGRRVGARMLGDMVAVSTCIAPVPHVIMQNSMRLPARPKTCYARPLVSRYADSGELI 622
571 QGOLGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENYFKVWDAADIQIASTFVEL 630
623 EGOLGDNELRLEQNPLECTVGHKRYFVFGDGYVYFEEYASHQVSRADVPVSTFVDL 682
631 NLTLEDREILPLSVYTKELRDVGVDYAEVARRNQLHELKFDYDINKVIEVDYNTVA--- 687
683 NLTLEDHEFLPEVYVTRQEIKDSGLLDYAEVQRNQMHALFSDIDHIINDTTNAALMD 742
688 GLQEF 692
743 GLFRF 747

RESULT 13
VGLB_HSV2S
ID VGLB_HSV2S STANDARD; PRT; 885 AA.
AC P24994;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 2 / strain SA8) (Simian agent 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_Taxid=103116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B264;
RA Borchers K., Weigelt W., Buhk H.-J., Ludwig H., Mankertz J.;
RT "Conserved domains of glycoprotein B (gB) of the monkey virus, simian
agent 8, identified by comparison with herpesvirus gBs.";
RL J. Gen. Virol. 72:2299-2304(1991).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GG, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X56935; CAA40236.1; -
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 885 GLYCOPROTEIN B.
FT DOMAIN 35 715 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 716 731 POTENTIAL.
FT TRANSMEM 737 756 POTENTIAL.
FT TRANSMEM 760 780 POTENTIAL.
FT DOMAIN 781 885 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 659 659 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 885 AA; 97811 MW; 39E495B329AB94E4 CRC64;

Query Match 30.5%; Score 1714.5; DB 1; Length 885;
Best Local Similarity 48.2%; Pred. No. 2,1e-79;
Matches 329; Conservative 108; Mismatches 210; Indels 35; Gaps 5;

QY 20 NSPSTQNTSVREVSSVOLSEESTFYLCPPVGVSTVIRLEPPKRCPPRKATENGEGI 79
DB 68 NASVEAGRATLREDLEIRKARDGATFYVCPPTGATVYQEQPRCPRAPDQNTQTEGI 127
QY 80 AILFKENISPYKFKVLYYKNIQTITWTGTYRTQITNRYTDTPTVYSIEITDLIDGKR 139
DB 128 AVFRENIAPIYKFKATMYKDYTSQVWEGHRYSQWGFEDRAPVPFEVMDKINAKGV 187
QY 140 CSKARYLNRYVVEAFDAGEKQVLLKPSKNTPESSRAWHNTTNYTVWGSFWIYRTG 199
DB 188 CRSTAKYVRNNMESTAFHRDDHSDMALPKAKAATRTSRGHTTDLKYNPARVEAFHYG 247
QY 200 TSVNCIIVEEMDARSVPYSPYFAMANGDIANSFPYGLSPPEAAEPGYQDNFKQLDSY 259
DB 248 TTVNCIIVEEARSVPYDFEVLATGDFYVMSPFYGR-DGSHGEHTAYAAADRFRQDGY 306
QY 260 FSDMDLDRKRAKSLPVKRNFLITSHFTVGDWAPKTRVCSMTKWKVEITMLRATVNGRY 319
DB 307 YERDLSTGRRAAPVTRNLTTPKFTVGDWAPKRRSVCTLNKREVDMLRAEYGPSPR 366
QY 320 FMARELSATFISNTTEFPDNRILGQCIKREAAEAIQIFRTKYNDSHVGVHVVQVFLAL 379
DB 367 FSSAALSTFTTANRTEYALSRLDADCVGREAREAVDRIFLRRYNGTHVKVGQVYLLAT 426
QY 380 GGFIVAYQVPLSKSLAHMYLRELNRDNTDEMLDLVNKHAIYKKNATSLSLRRDIRNA 439
DB 427 GGFLIAYQPLLSNALVELYVRELVE-----QTRRPAGD 461
QY 440 PNKRIT----LDDTTA--IKSTSSVQFAMLOFLYDHIQTHINDMFESRIATWCELONREL 493
DB 462 PGEATPGSDVPDPSPVERIKTSSVEFARLQFTYDHIQHVNEMLGRIAAWCOLQNEL 521
QY 494 VLWHEGKINPSATASATLGRRAAKMGDVAASCTAIDAESVTLQNSMRVITNTQTC 553
DB 522 TLWNEARRLNPCAIASATVGRRAARMGLGDMVAVSTCVPAVDNVIQNSIGVAAAPGTC 581
QY 554 YSRPLVLSYGENQNIQOLGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENYV 613
DB 582 YSRPLVLSRYEADGPLVEGQLGEDNEILRLDALEPCTVGHRRYTFGAGYVYFEEYAS 641
QY 614 KMWDAADIQIASTFVELNLTLEDREILPLSVYTKELRDVGVDYAEVARRNQLHELK 673
DB 642 HOLGRADVTVTSTFTNLNLTLEDHEFLPEVYVTRQEIKDSGLLDYAEVQRNQLHACRF 701

```

QY 674 YDINKVIEVDN---YAGLQEF 692
 DB 702 ADIDTVIKADAAHAAFLAGLYSF 723

RESULT 14
 VGLB_HSVB2
 ID VGLB_HSVB2 STANDARD; PRT; 917 AA.
 AC P12641;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B-1 precursor.
 OS Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89305231; PubMed=2841793;
 RA Hammerschmidt W., Contraths F., Mankertz J., Pauli G., Ludwig H.,
 RA Buhk H.-J.;
 RT "Conservation of a gene cluster including glycoprotein B in bovine
 RT herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";
 RL Virology 165:388-405(1988).
 RN [2]
 RP SEQUENCE OF 1-200 FROM N.A.
 RX MEDLINE=88306232; PubMed=245278;
 RA Hammerschmidt W., Contraths F., Mankertz J., Buhk H.-J., Pauli G.,
 RA Ludwig H.;
 RT "Common epitopes of glycoprotein B map within the major DNA-binding
 RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
 RT virus type 1 (HSV-1).";
 RL Virology 165:406-418(1988).
 CC -1- FUNCTION: GBL IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE
 CC PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A
 CC SYNCRITIAL PHENOTYPE.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M21628; AAA46053.1; -;
 DR EMBL; M21632; AAA46052.1; -;
 DR PIR; C29242; VGBEBH.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 917 GLYCOPROTEIN B-1.
 FT TRANSSEM 578 594 POTENTIAL.
 FT TRANSSEM 770 786 POTENTIAL.
 FT TRANSSEM 795 811 POTENTIAL.
 FT CARBOHYD 48 48
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 917 AA; 101882 MW; 1B96CBF50DB4D3F3 CRC64;
 Query Match 30.4%; Score 1710; DB 1; Length 917;
 Best Local Similarity 48.1%; Pred. No. 3.7e-79;

Matches 325; Conservative 117; Mismatches 202; Indels 32; Gaps 4;
 QY 20 NSSPSTQNTSREVSVSSQLSEESTFYLCPPPVGVSTVIRLEPPKCPPEPKATEWBGCI 79
 DB 110 NASEPADPAELRDLGLKSSDDPNFYVCPPTGATVVRLEEPKPELPGKLFNTEGI 169
 QY 80 AILFENISPYKFKVLYYKNIQTWTGTYTQITNRYTDRTPVSTEEITDLDGKGR 139
 DB 170 AVTEKENLAPYKFKATMYKAVTVASVWVGYSYNOQFNIFEDRAPIFEELVDRHGM 229
 QY 140 CSSKARYLRNNVYVEADRDAGEKQVLLKPSKFTPESRAWHHTTNETYTVWGSPIWRTG 199
 DB 230 CLSTAKYVRNNLETTAFHNDADHEMKLVPAESAPGLHRGHTTTLKNNPTGSAWIHRHG 289
 QY 200 TSVNCIVVEMDARSVPYSYFAMANGDIANSIPYGLSPPEAAAEPMGYPODNFKQDLSY 259
 DB 290 TVDCIVDEVEAKSSYPINEFVLATGDFVYASPFPGYR-DGSHSEHNAADREKQVDGF 348
 QY 260 FSDLDKRRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKWKVTEMLRATVNGRY 319
 DB 349 FPRDFGTGRRHGSPVTYNLLTTPMTVGNWAPKRPSCVMTKWREPEMLRAEYGSFR 408
 QY 320 FMARELSATFISNTTEFDPNRIILGQCICKREAAIEQIFRTKNDSHVKGHVQYFLAL 379
 DB 409 FTSNALSATFTNTLTQTSLSRVDLGDGCVGEAREADRIYLEKYNNTNHLRVGSVQYLLAT 468
 QY 380 GGFIVAYQPVLSKSLAHMYLRELNRKNTDEMILNVNKHAIYKKNATLSRLRRDIRNA 439
 DB 469 GGFIIAYQPLLNNLADLYVKELMREQL-----KPEERKLNATT----- 508
 QY 440 PNKRITLDDTATKSTSSVQFAMLOFLYDHIQTHINDMFSRIATAMCELQNLRLVWHEG 499
 DB 509 -----DGKVITTTSSVEFARLQFTYNHQKHNEMFGRMVAVSWCELQNLQELTLNEA 560
 QY 500 IKNPSTASATSLGRRVAANKMGLGDAVAVSCTAIDAEVSLQNSMRVITSTNTCYSRPLV 559
 DB 561 KKNPSSAIASTVTLHRRVSAACMLGDVLAISTCVAPAEVNIQNSMRIPSPGTCYSRPLL 620
 QY 560 LFSYGENQGIQGLQENNELPTLEAVEPCSANHRRYFLFGSGYALFENYVFKVYDAA 619
 DB 621 SFRVDSGEELMEQGLGENNEIRLDRDAVEPCSVGHKKYFLFGAGYVFEYTYSHQUSRS 680
 QY 620 DIQIASTFVEINLTLEDRILPLSVYTKELRDVGLDYAEVARRNQLHKLKFDYINKV 679
 DB 681 DITAVSTFDLNTLMLEDFEVPLEVTROEIKDGLLDYAEVQRRNQLHALRFADIDTV 740
 QY 680 IEVDTN---YAGLQEF 692
 DB 741 IKADPNNAIFAGLHGF 756

RESULT 15
 VGLB_HSV1F
 ID VGLB_HSV1F STANDARD; PRT; 903 AA.
 AC P06436;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL27.
 OS Herpes simplex virus (type 1 / strain F).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85083254; PubMed=2981343;
 RA Peillett P.E., Kousoulas K.G., Pereira L., Roizman B.;
 RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
 RT of monoclonal antibody-resistant mutants.";
 RT J. Virol. 53:243-253(1985).
 RN [2]

RP SEQUENCE OF 1-176 FROM N.A.
RA MEDLINE-88306232; PubMed-2457278;
RA Hamerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -I- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -I- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GG, GD, GI, AND GE.
CC -I- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14164; AAA45776.1; -;
DR EMBL; M21633; AAA45788.1; -;
DR PIR; A03750; VGBEB1.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 29
FT CHAIN 30 903 GLYCOPROTEIN B.
FT DOMAIN 31 729 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 730 745 POTENTIAL.
FT TRANSMEM 751 770 POTENTIAL.
FT TRANSMEM 774 794 POTENTIAL.
FT DOMAIN 795 903 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 903 AA; 100104 MW; 73BDCA7813DB35E8 CRC64;

Query Match 30.4%; Score 1705.5; DB 1; Length 903;
Best Local Similarity 47.6%; Pred. No. 6.1e-79;
Matches 326; Conservative 112; Mismatches 206; Indels 41; Gaps 6;

QY 18 GTNSPSTQNTISREVSVSVQLSEESTFYLCPPVGVSTVIRLEPPKCPKATWGE 77
DB 84 GDNATVAAGHATLREHLRDKAENTDANFYCCPPPTGATVQFEQPRCPTREGEQNTY 143
QY 78 GIALFKENISPKFKVLYKNIQTWTGTTVQITNRYTDRTPVSIETDLIDGK 137
DB 144 GIAVFKENIAPYKATKYKDVTSQVWFGHRYSGFMGIFEDRAPVPFEVIDKINAK 203
QY 138 GRCSKARYLRNNVVEAFDRDAGEKQVLLKPSKNTPESSRAWHHTTNETYTWGSPWYR 197
DB 204 GVCSTAKYVRNNLETTAFHRDDHETDMLKANAATRTSRGWHHTDLKYNPSRVEAFHR 263
QY 198 TGTSTNCIVEMDARSVPYFAMANGDIANISPFYGLSPPEAAAEPMGYPQDNFKOLD 257
DB 264 YGTTVNCIVEEDARSVPYDEFVLATGDFVYMSPFYGR-EGSHTEHTSYAADRKFQWD 322
QY 258 SYFSDMLDKRKASLPVKRNFLLTSHFTVGVWDAPKTRVCSMTKWKVEVTEMLRATVNGR 317
DB 323 GFYARDLTAKARAPTTRNLLTTPKFTVMDWVPKPSVCTMTKWQVDEMLRSEYGGG 382
QY 318 YRFMARELSATFISNTEFFDNPRIILGQICKREAAEAEQIFRTKYNDNSHKVGHVQYFL 377
DB 383 FRFSSDAISTFTTNLTETPLSRVDLGCIGKDARDAMDRIEARRYNATHIKVGPQYIL 442

QY 378 ALGGFIVAYQPVLSKSLAHMYLRELMDRNRDTEMLDLYNNKHAIYKKNATSLSRRLDRIR 437
DB 443 ANGGLIAYQPLLSNTLAELYRE-----HLREQSR 473
QY 438 NAPNRKITLDDTTA-----IKSTSSVOFAMLOFLYDHIQTHINDMFESRIATAWCELQNR 491
DB 474 KPPNTPPPPGCASANASVERIKTTSIEFARLOFTYNIQIRVNDMLGRVAIAWCELQNH 533
QY 492 ELVLWHEGIIKINPSATATLGRVAAKMLGDVAASVSSCTAIDAESVTLQNSMRVITSTN 551
DB 534 ELTLWNEARKLNPNIAISATVGRVYSARMGLGDVMAVSTCVPAADNVIVQNSMRISRRPG 593
QY 552 TCYSRPLVFSYGENQGN-IQOIGENNELLPLEAVPEPCSANHRRYFLFGSGYALFENY 610
DB 594 ACYSRPLVSFRY-EDQGPLVEGQGENNELRLTRDAIEPCTVGHRRYFTFGGGYVYFEY 652
QY 611 NFVKMVDAAIDQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVDVYAEVARNOLHE 670
DB 653 AYSHQLSRADITTVSTFFIDLNTMLEDHDFVPLEVYTRHEIKDSGLLDYTEVQRNQLHD 712
QY 671 LKFDYDINKVIEVDN---YAGLQEF 692
DB 713 LRFADIDTVIRHADANAAMFAGLGAF 737

Search completed: October 8, 2003, 17:03:40
Job time : 23.5547 secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:00 ; Search time 37.2016 Seconds
(without alignments)
2807.386 Million cell updates/sec

Title: US-09-147-052-4
Perfect score: 5619
Sequence: 1 MHYFRNCIFLLIVLYGTN.....SSNENADKIPGRRPCTFL 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3601	64.1	865	1	VBBERB		glycoprotein B pre
2	1898.5	33.8	948	2	A56602		glycoprotein B hom
3	1779.5	31.7	980	1	VBEC6		glycoprotein B pre
4	1775.5	31.6	868	1	VBEC31		glycoprotein B - h
5	1775	31.6	913	1	VBBERP		glycoprotein gli p
6	1761.5	31.3	979	2	JH0109		glycoprotein i4 pr
7	1760.5	31.3	919	1	VBEGH		glycoprotein B pre
8	1755.5	31.2	975	2	T42576		probable envelope
9	1743	31.0	943	2	B48474		glycoprotein B - f
10	1741.5	31.0	980	1	VBEB2H		glycoprotein B pre
11	1735	30.9	904	1	VBEB2		glycoprotein B pre
12	1731.5	30.8	932	1	VBEBEC		glycoprotein gi pr
13	1727.5	30.7	920	1	B48349		glycoprotein B pre
14	1714.5	30.5	885	1	VBESA		glycoprotein B pre
15	1714	30.5	904	1	VBEBK2		glycoprotein B pre
16	1710	30.4	917	1	VBEBH		glycoprotein B pre
17	1705.5	30.4	903	1	VBEB1		glycoprotein B pre
18	1705.5	30.4	904	1	VBEBW7		glycoprotein B pre
19	1699.5	30.2	903	1	VBEBK1		glycoprotein B pre
20	1671.5	29.7	933	1	B48349		glycoprotein B pre
21	1604	28.5	928	1	VBEBG		glycoprotein gi pr
22	1337.5	23.8	883	1	VBEBIS		glycoprotein B pre
23	1325	23.6	883	1	VBEBIL		glycoprotein B pre
24	1322	23.5	873	2	S26690		glycoprotein B - i
25	831.5	14.8	702	2	S48754		major surface prot
26	778.5	13.9	907	1	VBETE		glycoprotein B pre
27	760.5	13.5	854	2	T03107		glycoprotein B - a
28	756	13.5	906	1	VBEC1		glycoprotein B pre
29	754	13.4	702	2	S48753		major surface prot

30	750	13.3	830	1	A44047	glycoprotein B pre
31	745	13.3	830	2	T43999	glycoprotein B lim
32	744	13.2	650	2	S48751	major surface prot
33	743	13.2	649	2	S48752	major surface prot
34	741	13.2	822	2	T41941	glycoprotein B - h
35	740.5	13.2	874	2	S25530	glycoprotein B - b
36	740	13.2	830	1	B44047	glycoprotein B pre
37	740	13.2	830	2	T44186	probable glycoprot
38	738.5	13.1	831	1	VBEG6S	glycoprotein B - h
39	737	13.1	928	1	VBEMC	glycoprotein B pre
40	713	12.7	647	2	A49218	hemagglutinin homo
41	709.5	12.6	857	1	QOBEIL	glycoprotein B - h
42	700	12.5	807	2	T42924	glycoprotein B - a
43	695	12.4	808	1	VBESM	glycoprotein B pre
44	679.5	12.1	874	2	S55602	glycoprotein B - e
45	575.5	10.2	386	2	S48755	major surface prot

ALIGNMENTS

RESULT 1

VBBERB
glycoprotein B precursor - Marek's disease virus (strain RBLB)
C:Species: Marek's disease virus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A32402; B32402
R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.
J. Gen. Virol. 70, 1789-1804, 1989
A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog
A:Reference number: A32402; MUID:89293086; PMID:2544666
A:Accession: A32402
A:Molecule type: DNA
A:Residues: 1-865 <ROS>
A:Cross-references: GB:DI3713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837
A:Accession: B32402
A:Molecule type: Protein
A:Residues: 250-271;304-330 <ROS2>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-865/Product: glycoprotein B #status predicted <GPB>
F:709-728/Domain: transmembrane #status predicted <TN2>
F:732-752/Domain: transmembrane #status predicted <TN2>
F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match	64.1%	Score 3601;	DB 1;	Length 865;
Best Local Similarity	85.9%;	Pred. No. 4.5e-173;		
Matches 719;	Conservative 16;	Mismatches 60;	Indels 42;	Gaps 8;
Qy	1	MHYFRNCIFLLIVLYLXGTNSSPSTQNTVSREVSVSSVOLSEESTFYLCPPPVGVSTVRL	60	
Db	1	MHYFRNCIFLLIVLYLXGTNSSPSTQNTVSREVSVSSVOLSEESTFYLCPPPVGVSTVRL	60	
Qy	61	EPKRCPEPRKATWEGGIALFKENISPYKFTLYKNIQTTWTGTYRQITNRYT	120	
Db	61	EPKRCPEPRKATWEGGIALFKENISPYKFTLYKNIQTTWTGTYRQITNRYT	120	
Qy	121	DRTPVSTEEITDLIDGRCSSKARYLRNNVYVAFDRDAGEKQVLLKPSKFNTPESRAW	180	
Db	121	DRTPVSTEEITDLIDGRCSSKARYLRNNVYVAFDRDAGEKQVLLKPSKFNTPESRAW	180	
Qy	181	HTTNETYTVMGSPWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE	240	
Db	181	HTTNETYTVMGSPWIYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISFFYGLSPPE	240	
Qy	241	AAAPMGYPQDNFKQLDSYFSMDLDRKKASLPVKRNLITSHFTVGVNDWAPKTRVCSM	300	
Db	241	AAAPMGYPQDNFKQLDSYFSMDLDRKKASLPVKRNLITSHFTVGVNDWAPKTRVCSM	300	
Qy	301	TKKREVTMLRATNGRYRFRNARELSATFISNTTTEFPNRIILGCIKREAAAEIQIFR	360	
Db	301	TKKREVTMLRATNGRYRFRNARELSATFISNTTTEFPNRIILGCIKREAAAEIQIFR	360	

QY 361 TKYNDSHVKGHVQVFLAGGFIYVQVLSKSLAHMYLRELMDLNNKHA 420
 DB 361 TKYNDSHVKGHVQVFLAGGFIYVQVLSKSLAHMYLRELMDLNNKHA 420
 QY 421 LYKKNATSLRLRRDIRAPNKKITLDDTTALKSTSSVOFAMQLQFLYDHIQTHINDMFSR 480
 DB 421 LYKKNATSLRLRRDIRAPNKKITLDDTTALKSTSSVOFAMQLQFLYDHIQTHINDMFSR 480
 QY 481 TATACELQNLRELVLWHEGKINPNSATATLGRVAAKMGDVAVSSCTAIDAESVTL 540
 DB 481 TATACELQNLRELVLWHEGKINPNSATATLGRVAAKMGDVAVSSCTAIDAESVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLSYGNQNOGQGENNELPDLTLEAVPSCSNHRRYFLF 600
 DB 541 QNSMRVITSTNTCYSRPLVLSYGNQNOGQGENNELPDLTLEAVPSCSNHRRYFLF 600
 QY 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDA 660
 DB 601 GSGYALFENYFVKVMDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDA 660
 QY 661 EVARNQHLKFLYDINKVIEVDNYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
 DB 661 EVARNQHLKFLYDINKVIEVDNYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
 QY 717 ELTDLAKAMTLASLDQYAK-----IEASLSAYSEATVN-----NNLNA--- 758
 DB 717 ELTDLAKAMTLASLDQYAK-----IEASLSAYSEATVN-----NNLNA--- 758
 QY 709 VVGAAGAIVSTIGSVAFMSNPFALGALGLIAGLVAFLAYRYVKNLKSNNPKALYP 768
 DB 709 VVGAAGAIVSTIGSVAFMSNPFALGALGLIAGLVAFLAYRYVKNLKSNNPKALYP 768
 QY 759 -TLEOLKMAKNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQATNL 809
 DB 759 -TLEOLKMAKNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQATNL 809
 QY 769 MTEVLKQAQRELHGEESDOLERTSIDER--KLEEARMIKYMALVSAERHEKKL 823
 DB 769 MTEVLKQAQRELHGEESDOLERTSIDER--KLEEARMIKYMALVSAERHEKKL 823

RESULT 2

A56602
 glycoprotein B homolog precursor - feline herpesvirus 1
 C:Species: feline herpesvirus 1
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999
 R:Accession: A56602
 R:Maeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Niikura, M.; Kai, Arch. Virol. 127, 387-397, 1992
 A:Title: Identification and nucleotide sequence of a gene in feline herpesvirus type 1
 A:Reference number: A56602; PMID:93090104; PMID:1333759
 A:Accession: A56602
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-948 <MAE>
 A:Cross-references: GB:S49775; NID:g261094; PIDN:AAB24381.1; PID:g261095
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBI:120002, NCBI:120003)
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein

Query Match 33.8%; Score 1898.5; DB 2; Length 948;
 Best Local Similarity 44.6%; Pred No. 1.3e-87;
 Matches 383; Conservative 158; Mismatches 260; Indels 57; Gaps 12;
 QY 18 GTNSPSPSTQNTSREVSVSSVOLSEEE---STFYLCPPPPGVSTVIRLEPPPRKCPPRKAT 73
 DB 93 GTPPKPTTDDTMSDMREALRASQIEANGPSTFYVCPGSGTSTVYRLEPPPRACPDYKLGK 152
 QY 74 EMGEGIALFKENTSPYKFKVLYKYNIIQTTWTGTYRQITNRYTDTPTVPSIEITDL 133
 DB 153 NTEGIAVIFKENTAPYKFKVLYKYNIIQTTWTGTYRQITNRYTDTPTVPSIEITDL 212
 QY 134 IDGKRCSSKARYLNNVYVFAEDRAGEKQVLLKPSKENTPESRAWHTTNETYVWGPS 193
 DB 213 IDRRGCMUSKADYVRNNTQTFADDEDPRELPLKPSKENTPQSRGWHTTNETYVWGPS 272
 QY 194 WYRTGTGTVNCIVEMDARSFPYFAMANGDIANISFPYGLSPPEAAAEPMGPQDNF 253
 DB 273 GFHHSSTVNCIVEMDARSFPYFAMANGDIANISFPYGLSPPEAAAEPMGPQDNF 331

QY 254 KOLDYSFQMDLDRKRKASLPVKRNFLLTSHFTVGDWAPKTRVCSMTKKWKEVTEMLRAT 313
 DB 332 QOIEGYVPLDLDRLQLGAPVSRNLEPHVTVANWTPKCGRVCTLAKWEIDEMLRDE 391
 QY 314 VNGRYRFAWELSATFISNTTFEDPNRILGOCIKREAAEQIFRKYNDSHVKGHV 373
 DB 332 YGGSYRFTVKTISATFISNTSOFENIRLGLCATKEAAEADRIYKSKYTHIQGTGL 451
 QY 374 QVFLAGGFIYVQVLSKSLAHMYLRELMDLNNKHAHYKKNATSL--SR 431
 DB 452 EYTLARGEFLAFRPMISNELAKLYINELARSNRVVDLSALLNPSETVORTRGVPSNQ 511
 QY 432 LRDRIRNAPNKKI-TLDDTTAKSTSSVOFAMQLQFLYDHIQTHINDMFSRIATAMCELQ 490
 DB 512 HRSRSTTEGGIETVNNASLLKTTSSVEFAMIQPAYDIQAHVNEMLSRATAMCTLQN 571
 QY 491 RELVLWHEGKINPNSATATLGRVAAKMGDVAVSSCTAIDAESVTLQNSMRVITST 550
 DB 572 REHVLTETLKLNPGGVSMALERRVSARLLGDVAVTQCVNISSGHVYIQNSMRVTGSS 631
 QY 551 WTCYSRPLVLSYGNQNOGQGENNELPDLTLEAVPSCSNHRRYFLFGSGYALFENY 610
 DB 632 TTCYSRPLVSPALNDSEYIEGQGENNDLIVERLIEPCTVNNKRYFKFGADYVYFDEY 691
 QY 611 NFVKVMDAADIQIASTFVELNLTLLEDEILPLSVYTKKEELRDVGVLDAEVARNQHL 670
 DB 692 AYVRKVPLEISELISAYVDNLTLLEDEILPLSVYTKKEELRDVGVLDAEVARNQHL 751
 QY 671 LKFDYDINKVIEVDNYA-----YAGLOEFGC-----MSITKKDANPN 706
 DB 752 LKFDYDINKVIEVDNYA-----YAGLOEFGC-----MSITKKDANPN 706
 QY 707 GOTOLEAARMELTDLINA-----KAM---TLASLDQYAKIEASLSAYSE 748
 DB 812 PFGALVGLLILAGIVAAFLAYRISRLRANPKALYPVTTNQLQAKSPASTAGGSD 871
 QY 749 AETVNNLNATLEQLKMAK-TNLESAINQANTDKTTFDNEHPNLVEA---KALKT--TL 802
 DB 872 PCVDFDEEKLMOAREMIKYSVLSVAMEQ-QBHKMKKNKGPAILTSLTNMALRRRGPK 930
 QY 803 EQRATNLEGLSTAYNQI 820
 DB 931 YORLNLDSGGDDTETNLV 948

RESULT 3

VBEC6
 glycoprotein B precursor - equine herpesvirus 1 (strain Ab4p)
 C:Species: equine herpesvirus 1
 A:Note: host Equus caballus (domestic horse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
 C:Accession: G36798
 R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J. submitted to GenBank, March 1992
 A:Description: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A36805
 A:Accession: G36798
 A:Molecule type: DNA
 A:Residues: 1-980 <TEL>
 A:Cross-references: GB:M86664; NID:g330791; PIDN:AAB02468.1; PID:g330825
 R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J. Virology 189, 304-316, 1992
 A:Title: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A41831; PMID:92295566; PMID:1318606
 A:Contents: annotation; possible protein-coding frames
 A:Note: neither amino acid nor nucleotide sequence is given
 C:Genetics:
 C:Gene: 33
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-85/Domain: signal sequence #status predicted <SIG>
 F:67-78/Domain: transmembrane #status predicted <TM1>

F:86-980/Product: glycoprotein B #status predicted <MAT>
F:850-868/Domain: transmembrane #status predicted <TM2>
F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (cova

Query Match 31.7%; Score 1779.5; DB 1; Length 980;
Best Local Similarity 41.9%; Pred. No. 1.3e-81;
Matches 379; Conservative 132; Mismatches 260; Indels 133; Gaps 17;

QY 19 TNSPSTONVTS-----REVSSVOLSEESTFYLCPPPVGST 56
DB 88 TTPSPPTSTPTSMTHSGTVDPTLLPTDPLRLAVRESGILAEDEGDFYTCPPPTGST 147
QY 57 VRLPPKPCPRKATWEGEGIALFKENISFYKLYKNIQIITWTGTYRQIT 116
DB 148 VVRIEPPRPKFDLGRNFTGEGIAFYKFNIAPIYKFRANVYKDIIVTVRWGYSHTSL 207
QY 117 NRYTORTVSIIEITLDLIDGKRCSSKARYLNNVYVEAFDAGEKQVLLKPSKFNTP 176
DB 208 DRYNDVPVSVIEIFGLIDSKCKSKARYLNNVYVEAFDAGEKQVLLKPSKFNTP 267
QY 177 SRWHTTNET--YTVWSPWYRTGTSVNCIVVEEDARSFVYSYFAMANGDIANISPPY 234
DB 268 ARAWOTNTDTSYGV--MPWRIYTSVNCIVVEEDARSFVYSYFAMANGDIANISPPY 326
QY 235 GLSPPEAAA--BPMGYPDQNFQSDYFSDMDLDRKASLPVKRNLITSHFTVGVWDAP 292
DB 327 GL---RAAARIENHSYAQRFRQVEGRPRDLSKLAQAEPTKMFITPHVTVSNWTE 383
QY 293 KTRVCSMTKWEVTEMLATNGRYRFRMARELSATFISNTTFDPNRIILGQCCKRAE 352
DB 384 KQVEACTLKWEVDELVRFGSVFTIRISSTFISNTQFKLESAPLTCYSKEAK 443
QY 353 AAIEQIFRTKYNDSHVGVHVOYFALGDFIVAYOPVLKSLAHMYLRELMRDNRDTEML 412
DB 444 EADISYKQYESTHVSFGDVEYILARGGFLAFRPMLSNELARLYLNELVSNRTYDLK 503
QY 413 DLVNNKHAIYKKNATSLSLRLDIRNAPNRKTLD-----DTTA----- 451
DB 504 NLLNPN-----ANNNTTTRRRSLLSVPEPQDGVHREQILLRLHKAVEATATGDS 559
QY 452 -----IKSTSSVQFAMLOFLDHIQTHNDMFRIATATWCELQNLRELVLWHEGKI 502
DB 560 NVTAKOLELKITSSIEFAMLOFAYDHIQSHVNEMLSRITATWCTLQNKERTLWNEVKI 619
QY 503 NPSATASATLGRVRAAKMLGDVAVSSCTAIDAESVTYQNSMRVITSTWTCYSRPLVLS 562
DB 620 NPSAIVSATLDERVAARVLGDVIAITHCAKIEG--NYVLQNSMRSDS--NTCYSRPPVTF 677
QY 563 Y---GENOGNIQOLGKNNELLPTLEAVPCSANHRRYELFGSGYALFENYFVKMVDAA 619
DB 678 IYKNANRSGIEQLGEEIEETERKLEPCALNOKRYKFEYVYENITFTVRKVPPT 737
QY 620 DIQIASTFVELNLTLEDREILPLSYTKEELRGVLDYAEVARRNQHLKFDYDINKV 679
DB 738 EIEVISTYVELNLTLEDREILPLSYTKEELRGVLDYAEVARRNQHLKFDYDINKV 797
QY 680 IEVDIN-----YAGLQEG-----CMSTTKDANPNQGT--- 709
DB 798 VVVDNTAVIMQIASFFKGLGVGEAVGTLVGAAGAVVSTVSGIASFLNPNFGLAIGL 857
QY 710 -----QLEAARMELTDLINAKMTLASIQDYAKIEASLSAYSEATV 752
DB 858 LVIAGLVAFFAYRYVMQIRSNPMKALYPTITKALKNKAKTSYGONEEDGSDFEAKLE 917
QY 753 N-----NLLNATLEQLKMA--KTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803
DB 918 EAREMKYMSVSALEKQEKKAIKKNSGVGLIASNVSKLALRRPGKY-----TRLQ 969
QY 804 QRAT 807
DB 970 QNDT 973

RESULT 4

VGBC31

glycoprotein B - human herpesvirus 3
N:Alternate names: glycoprotein II
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: E27214
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: E27214
A:Molecule type: DNA
A:Residues: 1-868 <D>
A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27914.1; PID:g60020
C:Genetics:
A:Gene: 31
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:697-713/Domain: transmembrane #status predicted <TM1>
F:726-742/Domain: transmembrane #status predicted <TM2>
F:84,194,372,416,440,454,557,623,781,836,864/Binding site: carbohydrate (Asn) (cova

Query Match 31.6%; Score 1775.5; DB 1; Length 868;
Best Local Similarity 50.3%; Pred. No. 1.7e-81;
Matches 335; Conservative 122; Mismatches 186; Indels 23; Gaps 7;

QY 31 REVSSVOLSEESTFYLCPPPVGVSTVIRLEPPKPCPRKATWEGEGIALFKENISPY 90
DB 41 REAIHKSQDAETKPTFYVCPPTGSTVIRLEPRTCPDYHLGKNFTGEGIAVYKENTAA 100
QY 91 KKVLYLYKNIQIITWTGTYRQITNRYTDRTPVSEITDLIDGKRCSSKARYLRNN 150
DB 101 KFKATYVXDVAVSTAWAGSSYQITNRYADRVIPVSEITDIDFGKSSKATYVRN 160
QY 151 VVVEAFDRDAGKQVLLKPSKNTPESSRAWHHTNRYTVWGSPIYRTGTSVNCI 210
DB 161 HKVEAFNEKNDQDMLPIASKYNSVGSKAWHHTNDYIMVAGTPTTGTGTSVNCI 220
QY 211 ARSVFYSYFAMANGDIANISPFYGLSPPEAAEPGYPQDNFKQLDSYFSMDLDRKRA 270
DB 221 ARSIPYDPSGLSTGDIYMSPPFGLR--DGAYREHSNYANDRPHQEGYRQRLDTRALL 279
QY 271 SLPVKRNFLITSHFTVGVWDAPKTRVCSMTKWEVTEMLATVNGRYRPMARELSATFI 330
DB 280 E-PAARNFLVTPHLTVGNWPKRTEVCSLVKWEVEDVYRDEYAHNFRFTKTLSTTFI 338
QY 331 SNTREDPNRIILGQCCKRAEAAIEQIFRTKYNDSHVGVHVOYFALGDFIVAYOPVL 390
DB 339 SETNEFNLIHLSQCKEERAIINRIYTRTNSSSHVRIGDIQTLARGGFVVFQPLL 398
QY 391 SKSLAHMYLRELMRDNRDTEMLDVNNKHAIYKKNATSLSLRLDIRNAPNRKITDDTT 450
DB 399 SNSLARLYLQELVRE-----NTNHS--PKHPTNTRSRSV-----PVELRANR 441
QY 451 ATKSTSSVQFAMLOFLDHIQTHNDMFRIATATWCELQNLRELVLWHEGKIINPSATASA 510
DB 442 TITTTSSVFAFAMLOFTYDHIQEHVNEMLARISSSWQQLQNRERALSGLFIPNSALST 501
QY 511 TLGRRVAAMLDGVAVSSCTAIDAES--VTLQNSMRVITSTNCTYSRPLVLSYGENQGN 569
DB 502 ILDQVRKARILGDVIVSNCPGLSDTRIILQNSMRVSGSTTRTCYSRPLISIVSLNGSGT 561
QY 570 IQQGLKNNELLPTLEAVEPCSANHRRYELFGSGYALFENYFVKMVDAAIDQIASTFVE 629
DB 562 VEGQLGTDELINSLRDLLEPCVANHRYFLFGHHYVYEDYRVYRIAVHDVGMISTYYVD 621
QY 630 LNLTLLEDREILPLSYTKEELRGVLDYAEVARRNQHLKFDYDINKVIEVDNYA-- 687
DB 622 LNLTLKREEMFLQVYTRDELRTGLDYSIQQRNQMSLRFYDIDKVVQVYDSTAIM 681
QY 688 -GLQEF 692
DB 1: : :

Db 368 TTPHVTWNNWPEKKIEACTLTWKKEVDLVRDEPRGSGYRFTIRSISSFTFISNTTQFKLE 427
 QY 340 RIILGQCIRKRAEAAEIOIFRKYNDYSHVKGHVYFALGQFIVAYOPVLKSLAHMYL 399
 Db 428 DAPLTDVCSKEAKDAISYRKQYESTHVFSGDVEFYLAGRGFLIAFRPMISNELARLYL 487
 QY 400 RELMRDNRDMDLQVNNKHATYKKNAT'S-----LSRLRDRN 438
 Db 488 NELVRSNTYDLKNLL-NNANHTNRTRRSLLSTPEPTQESLHREQIILHRKRAVE 546
 QY 439 APN-----RKITLDDTTAKTSSVQFAMLOFLYDHIQTHINDMFESRIATAMCELOREL 494
 Db 547 AANSTNSSNVTAQLELLIKTASSIEFAMLOQAYDHIQSHVNEMLSRATFANCTLQNKERT 606
 QY 495 LWHEGILKINPSTATGRVAAKMGDVAVSSCTAIDAESTVLQNSMRVITSTNTCY 554
 Db 607 LWNEMKVPNPAISATLDERVAARVLGDVIAITHCVKIEG-NVYLQNSMR-SSDSNTCY 664
 QY 555 SRPLVLFSGYGN---QGNIOQOLGNNELLTLEAVEPCSAHRRYFLFGSGYALFENYN 611
 Db 665 SRPPVFTITNANSRGITIEQOLGEENEVYTERKLIIEPCAINQRYFKGKEYVYIENY 724
 QY 612 FVKVDAADIOIATFVELNLTLEDREILPLSVYTKELRDVGLVDAEVARRNQHL 671
 Db 725 YVRKVPPEIEIVISTYVELNLTLEDREILPLSVYTKELRDVGLVDAEVARRNQHL 784
 QY 672 KYDINKVIEVD-----TNVAGLOEFGCHSIKKDANPNNGQTOLEAARMELTDL 721
 Db 785 RFYDIDSVVNDTAVYMOGTATFFKGLRGV-----EAVGTLVGA 826
 QY 722 INAKMTLASIQDYAK-----IEASLSAYSEAEV-----NNLNATLEQLKMA 766
 Db 827 AGAVVSTVGIASFNNPFGGLAIGLLVAGLVAFFAYRYVMQLRSPMKALYPIITRS 886
 QY 767 KTNLESAINQANTDKTTFDNEHPNVEA-----YKALKTTLEQR 805
 Db 887 LKNKAKASYGQNDODDTSDFDEAKLEAREMIKYMVSVALEKQ 930

RESULT 9
 B48474
 glycoprotein B - feline herpesvirus 1 (fragment)
 C:Species: feline herpesvirus 1
 C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
 C:Accession: B48474
 R:Spatz, S.J.; Maes, R.K.
 A:Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B and a
 A:Reference number: A48474; MUID:94025559; PMID:8212548
 A:Contents: C-27
 A:Accession: B48474
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-943 <SPA>
 A:Cross-references: GB:S66371; NID:g435792; PIDN:AAB28559.1; PID:g435794
 A:Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBI:P138807)
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein

Query Match 31.08; Score 1743; DB 2; Length 943;
 Best Local Similarity 40.38; Pred. No. 8.3e-80;
 Matches 377; Conservative 161; Mismatches 267; Indels 130; Gaps 21;
 QY 3 YFRRNCIF-----FLIVLY-----GINSSP 23
 Db 22 YFQRCPFFSLGIAATGSHGNGSGLRLARYVSFIWLVLPVGRPVGSGSTSEQ 81
 QY 24 STQNTVSREV-----VSSVOLSEEE-----STFYLCPVPVGVSVIRLEPP 63
 Db 82 PRETVATPEGVHHQLOQIPPCRYEEALRASQIANGPSTYMCPPSPGSGIVRLEPP 141
 QY 64 RCPERKATEGEGTALFKENISPKYKVTLYNNIQTWTWTTVYQITNRYDRT 123

Db 142 RACPDYKIGKNFTGEGIAVIFKENIAPYKFKANIYKNIIMTWTWSSGSSYAVTNNRYTDRV 201
 QY 124 PVSIEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSFENTPESAWHTT 183
 Db 202 PVKQVEITDLIDRRGKCLSKADYVRNNYQFTAFDRDEDPRELPLKPPSTLSRVRGWH-T 260
 QY 184 NETYVWVSPWYIRFTGTSVNCIVEEMDARSVPFYSFAMANGDIANIPFYGLSPPEAAA 243
 Db 261 NETYKIVLLDFHHSGTSVNCIVEEDARSVPYDSFAISGTDVTHMSPFFGLR-DGAHV 319
 QY 244 EPMGYPDQNFQOLDSYFSMDLDRKASLPVKNRNLITSHFTVGVWDWAPKTRVCSMTKW 303
 Db 320 EHTSYSSRFOQIEGYPIIDLDTDTYGA-PVSRNFLETPTPHVTVANNWTPKSGRVCTLAKW 378
 QY 304 KEVTEMLRATVNGRYREMAELSATFISNTTEFDNRIILGOCIKREAEALQIFRKY 363
 Db 379 REIDEMPLMNI-GSYRFTAKTISATFISNTSQFELNRIUGDCATKEAAEADIRYKSY 437
 QY 364 NDHVKVGHVOYFIALGQFIVAYOPVLKSLAHMYLRELMRDNRTDEMCLVNNKHATYK 423
 Db 438 SKTHIQGTLETYLARGGFLIAFRPMISNELAKLYINELARSNT--VVDL-----SALLN 491
 QY 424 KNATSLSLRLRDIIRNPNRKI-----TLDDTTAKTSSVQFAMLOFLYDHIQTH 473
 Db 492 PSGETVORTKRSVPSNQHRSSRTIEGGIETVNNASLLKTTSSVEFAMLOFAYDIYQAH 551
 QY 474 INDMFSRIATACELQNLRELVLWHEGINKINPSTATATLGRVAAKMGDVAVSSCTAI 533
 Db 552 VNEMLSRATACWCTLQNRHVLWTETLKNPQGVVSMALERRVARSLLGDAVAVTQCVNI 611
 QY 534 DAESVTLQNSMRVTSNTCYSRPLVFSYGENGNTQOGLGENNELPTLEAVEPCSAN 593
 Db 612 SSGHYVQNSMRVTSNTCYSRPLVFSYGENGNTQOGLGENNELPTLEAVEPCSAN 593
 QY 594 HRRYFLFGSGYALFENYKVMVDAADIOIATSTVELNLTLEDREILPLSVYTKELRD 653
 Db 672 NKRYKFGADYVYEDYAYVKVPLSEIEELISAY-IRKSTLLEDREILH-SSYTRAELED 729
 QY 654 VGVDYAEVARRNQLHELKFDYDINKVIEVDTN-----YAGLOEFGC----- 694
 Db 730 TGPFDYSEIQRRNQLHALKFYDIDSIVRVNDNLVIMRGMANFFQGLGDVGAGFKVVLGA 789
 QY 695 -----MSITKNDANPNNGQTOLEAARMELTDLNA-----KAM---TLAS 731
 Db 790 ASAVISTVGVSSFLNPNPFGALAVGLLILAGIVAAFLAYRYSRLRANPKALYPVVTRN 849
 QY 732 LDYQATIASLSSAYSEAEVNNLNATLEOLKMAK-TNLESAINQANTDKTTFDNEHPN 790
 Db 850 LKQTAKPASTAGDSPPGVDDDEEKLMAQREMIKYMVSAMEQ-OEHKAMKKNKGA 908
 QY 791 LVEAY---KALKT--TLEQRATNLEGLSSTAYNQI 820
 Db 909 ILTSHLTNMLARRRGPYQRLNLDSDGDDTETNLV 943

RESULT 10
 VGBE2H
 glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)
 C:Species: equine herpesvirus 1
 C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
 C:Accession: A31241; J00008
 R:Whalley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Bell, C.W.; Woodworth, L
 J. Gen. Virol. 70, 383-394, 1989
 A:Title: Identification and nucleotide sequence of a gene in equine herpesvirus 1 and
 A:Reference number: A31241; MUID:89279217; PMID:2543744
 A:Accession: A31241
 A:Molecule type: DNA
 A:Residues: 1-980 <WHA>
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-85/Domain: signal sequence #status predicted <SIG>
 F:70-78/Domain: transmembrane #status predicted <TM>
 F:86-980/Product: glycoprotein B #status predicted <MAT>

F:853-868/Domain: transmembrane #status predicted <TM2>
F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (cova

Query Match 31.0%; Score 1741.5; DB 1; Length 980;
Best Local Similarity 41.4%; Pred. No. 1e-79;
Matches 375; Conservative 133; Mismatches 262; Indels 135; Gaps 19;

QY 19 TNSPSTQNTS-----REVSSVOLSEESTFYLCPPVGGST 56
DB 88 TTPSPSTPTSMTHSGTVDPTLPTPTDPLRLAVRESGILAEADGDFYCTPPPTGST 147
QY 57 VRLPEPRKCPPEKATGEGTALFLKENISPKVLYXKNIQTWTGTTVQRLT 116
DB 148 VVRIEPRCPFDLGRNTEGLIAVIFKENIAPKFRANYYKDIVVTRVWGYSYTSLS 207
QY 117 NRYDTRPVSYIEITDLIDGKRCSSKARYLRNNVYVEAFDRADAGEKV-LLKPSKFNTP 175
DB 208 DRYNDRVPVSVIEIFGLIDSKGCSKAEYLRDNIMHAYHDEDEVELDLCRPS-LQLR 266
QY 176 ESRAWHTTNET--YTVWGSPWYIRGTSVNCIVEEMDARSVPYSYFAMANGDIANSFP 233
DB 267 GARAWQTTNDTTSYVGW-MPWRYHTSTVNCIVEEYARSVPYDSFALSTGDIVYASPF 325
QY 234 YGLSPPEAAA--EPMGYPDNFKQLDSYFSDMLDKRKASLPVKRNFLLITSHFTVGDWA 291
DB 326 YGL---RAARIEHNSYAGERQVGYRPRDLDSKLQAEPTKFNFTPHVTVSWNT 382
QY 292 PKTRVCSMTKKWEVEMLRATVNGRYRFMARLSATFISNTFEDPNRIILGQCICKREA 351
DB 383 EKKVEACTLTWKKEVDELVRDFRGSYRFTIRSISSTFISNTTQFKLESAPLTECVSKEA 442
QY 352 EAAIEQIFPKYKNDHVHVQVFLALGGFIVAPQVLSLAHMYLRDLNRDNTDEM 411
DB 443 KEAIDSYYKKQVESTHFVSGDVEYLLARGGLFAFRPMLSNELARLYLNELVRSNRYDL 502
QY 412 LDLVNNKHAIYKKNATSLSLRDLIRNAPNRKLTLD-----DTTA--- 451
DB 503 KULLNPN-----ANNNTTTRRRSLSVPEPQTQGVHREQLILRLHRAVEATAGTDS 558
QY 452 -----IKSTSSVQFAMLOFLYDHIOTHINDMFSRIATAWCELONRELVLWHEGIK 501
DB 559 SNVTAQKLELIKTSSIEPAMLOFAYDHIQSHVNEMLSRATAWCTLQNKERTLWEMVK 618
QY 502 INPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTYLONSMRVITSNTCYSRPLVLF 561
DB 619 INPSATVATLDERVAARVLGDVIAITHCAKIEG-NVYLONSMRSDS-NTCYSRPPVTF 676
QY 562 SY----GENOGTQQLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKWDA 618
DB 677 TITKNANRSGSIEGQLGEENEIFTERKLIPECALNQRYFKGKEVYIYENTFFVRKVP 736
QY 619 ADIQIATFVELNLTLLEDEILPLSVYTKKEELRDVGVLDAEVARRNQLHELKFDYDINK 678
DB 737 TEIEVISTVELNLTLLEDEILPLSVYTRAELEDTGLLDYSIEQRNQLHALRFDYIDS 796
QY 679 VLEVDPN-----YAGLOEFG-----CMSITKKDANPNQGT--- 709
DB 797 VNVNDTAVIMOGIASFFGLGKVGGEAVTLVLGAAGAVSTVSGFASFLNPFGLATG 856
QY 710 -----QLEARMELTDLINAKMTLASLODYAKIEASLSAYSEAT 751
DB 857 LLVIAGLVAAPFAYRYVMQIRSNPKALYPIITKALKNNKAKTSYGQNEEDGSDFEAKL 916
QY 752 VN-----NNLNATLEQLKMA-KTNLESAINQANTKTFTDNEHPNLVEAYKALKTTTL 802
DB 917 EEAREMIKVMYSVALEKQEKAKKNSGVGLIASNVSKLARLRRRGPKY-----TRL 968
QY 803 EQBAT 807
DB 969 QQNDT 973

RESULT 11

VGEBB2
glycoprotein B precursor - human herpesvirus 2 (strain 333)
C:Species: human herpesvirus 2
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A26790; A45928
R:Stuve, L.L.; Brown-Shimer, S.; Pachl, C.; Najarian, R.; Dina, D.; Burke, R.L.
J. Virol. 61, 326-335, 1987
A:Title: Structure and expression of the herpes simplex virus type 2 glycoprotein gB
A:Reference number: A26790; MUID:87112925; PMID:3027364
A:Accession: A26790
A:Molecule type: DNA
A:Residues: 1-904 <STU>
A:Cross-references: GB:M15118; NID:g330256; PIDN:AAA45837.1; PID:g330257
Can. J. Microbiol. 33, 879-887, 1987
A:Title: The nucleotide sequence of herpes simplex virus type 2 (333) glycoprotein gB
A:Reference number: A45928; MUID:88079667; PMID:2446730
A:Accession: A45928
A:Molecule type: DNA
A:Residues: 1-34, 'AMPTV', 42-307, 'T', 309-481, 'R', 483-609, 'M', 611-664, 'R', 666-904 <2WA>
A:Cross-references: GB:M24771; NID:g341245; PIDN:AAA60540.1; PID:g623400
A:Note: the authors translated the codon ATG for residue 610 as Ile
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAT>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
F:772-792/Domain: transmembrane #status predicted <TM3>
F:82,136,250,393,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status
F:111-570,128-526,202-266,359-407,593-630/disulfide bonds: #status predicted

Query Match 30.9%; Score 1735; DB 1; Length 904;
Best Local Similarity 42.9%; Pred. No. 2e-79;
Matches 362; Conservative 143; Mismatches 269; Indels 70; Gaps 16;

QY 20 NSSPSTQNTSREVSSVOLSEESTFYLCPPVGGSTVIRLEPRKCPPEKATGEGI 79
DB 82 NATVAAGATLRAHLREIKVENADAQFYVCPPTGATVYVQFEQPRCPTRPEQNTTEGI 141
QY 80 AILFKENISPKVLYXKNIQTWTGTTVQRLTNRDNTVTSIEITDLIDGKR 139
DB 142 AVFKENIAPKFRANYYKDIVVTRVWGYSYTSLS-NTCYSRPLVLF 201
QY 140 CSKARYLRNNVYVEAFDRADAGEKVLLKPSKFNTPESRAWHHTNTTYTVWGSPWYIRTG 199
DB 202 CRSTAKYVRNNMETTAFHRDDHETDMELKPAKVATRTSGWHTTDLKYNPSRVEAFHRYG 261
QY 200 TSVNCIVEEMDARSVPYSYFAMANGDIANSIPFYGLSPPEAAAEPMGYPDNFKQLDSY 259
DB 262 TVVNCIVEEDARSVPYDFEVLATGDFVYMSPFYGR-EGSHTSHTSYAADRFGVDGF 320
QY 260 FSDMLDKRKASLPVKRNFLLITSHFTVGDWAPKTRVCSMTKKWEVEMLRATVNGRYR 319
DB 321 YARDLTTKARATSPTRNLLTTPKFTVANDVPRKAVCTMTKQVEDEMLRAEYGGSR 380
QY 320 FWARELSATFISNTFEDPNRIILGQCICKREAEEAIEQIFRTKYNDSHVHGVQVFLAL 379
DB 381 FSDAISTFTTTNLTQYSLSRVLDGICIGRDAREADRMFAKYNATHIKVGPQVYLAT 440
QY 380 GGFIVAYQVLSLAHMYLRDLNRDNTDEMDDLNNKHAIYKKNATSLSRRLDIRNA 439
DB 441 GGLIAYQLLSNTSLAEYLYREYMRQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNKRIKTDTTAKTSSVQFAMLOFLYDHIOTHINDMFSRIATAWCELONRELVLWHEG 499
DB 483 PSANASVE---RIKTTSSIEFARLQTYNHIQHVNDMLGRIAVANCELONHELTWNEA 539
QY 500 IKINPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTYLONSMRVITSNTCYSRPLV 559
DB 540 RKLNPNAIASATVGRVRSARMGLDVMAVSTCPVAPDNVIVQNSMRVSRPPGTCYSRPLV 599
QY 560 LFSYGENQGN-IOGQLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVKWDA 618


```

Db      600  SFRV-EDQGLIEGLGNNELRLTRDALEPCTVGHRRYFFGGYVVEEYASHQLSR 658
QY      619  ADIQAISTFVNLTLLEDREILPLSVYTKKEELRDVGVDYAEAVARRNQLHELKPYDINK 678
Db      659  ADVTTSTFDLNTIMLEDHFVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDT 718
QY      679  VIEWDTN---YAGLOEE-----CGMSITTKDANPNNGCOTLEARMELTDLI 722
Db      719  VIRADANAMPAGLCAFFEGMGDGLRAVGKVMVGVGVSVAVSVSF-----MSNPF 772
QY      723  NAKAMTLASLODYAKIESLASLSAYSEAEV-----NNLNLA-----TLEOLKMAKTNLESAI 774
Db      773  GALAVGLLVL-----AGLVAEFAPFYVLQLRNPMKALYPLTTKELKTSDFGGVGE 825
QY      775  NQANDTKTTFNEHPLNVEA-----YKALKTTLE--QRATNLEGLSSPAYNQIRNNLDL 827
Db      826  GEEGAEGGGFD--EAKLAAREMIRYMALVSAMERTHKARKKGTSSALLSSKVTNMVLK 883
QY      828  YNKA 831
Db      884  RNKA 887

```

RESULT 12

VGBEBC

glycoprotein gI precursor - bovine herpesvirus 1 (strain Colorado-1[Cooper-1])
 N:Alternate names: glycoprotein 11a; glycoprotein 16; glycoprotein g130; glycoprotein gV
 N:Contains: glycoprotein g1b; glycoprotein g1c
 C:Species: bovine herpesvirus 1
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
 C:Accession: A28877; A43380
 R:Whitbeck, J.C.; Bello, L.J.; Lawrence, W.C.
 J. Virol. 62, 3319-3327, 1988
 A:Title: Comparison of the bovine herpesvirus 1 gI gene and the herpes simplex virus type 1 gC gene
 A:Reference number: A28877; MUID:88300884; PMID:2841484
 A:Accession: A28877
 A:Molecule type: DNA
 A:Residues: 1-932 <WH>
 A:Cross-references: GB:M21474; NID:G330756; PIDN:AAA6055.1; PID:G330757
 R:van Drunen Little-van den Hurk, S.; Parker, M.D.; Fitzpatrick, D.R.; van den Hurk, J. Virology 190, 378-392, 1992
 A:Title: Structural, functional, and immunological characterization of bovine herpesvirus 1 glycoprotein gI
 A:Reference number: A43380; MUID:92410615; PMID:1326809
 A:Accession: A43380
 A:Molecule type: protein
 A:Residues: 68-87:505-516 <VAN>
 C:Superfamily: herpesvirus glycoprotein B
 C:Keywords: glycoprotein; transmembrane protein
 F:1-67/Domain: signal sequence #status predicted <SIG>
 F:68-932/Product: glycoprotein gI #status predicted <GIG>
 F:505-932/Product: glycoprotein g1b #status predicted <GIB>
 F:759-827/Product: glycoprotein g1c #status predicted <GIC>
 F:759-827/Domain: transmembrane #status predicted <TMN>
 F:105,153,441,483,640,706/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match      30.8%; Score 1731.5; DB 1; Length 932;
Best Local Similarity 49.8%; Pred. No. 3.1e-79;
Matches 341; Conservative 105; Mismatches 218; Indels 21; Gaps 7;

QY      18  GTNSSPSTQNTSREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPRPKCPPEPRKATEWGE 77
Db      97  GDDAASPDSNTDVRALRLAQAAGENSRRFFVCPSPGATVVRAPRCPPEYGLGRNTE 156

QY      78  GIALPKENTSPYKFKVLYKNIQTWTGTYTQIRYDTRTPVSEETDLDLIGK 137
Db      157  GIGVYKENTAPYTFKAYI-YKNVIVTTWAGSTYAAITNQYTDVRVPVGMGEITDLDVKK 215

QY      138  GRCSKARYLRNNVVEAFDRDAGEKOVLLKPSKENTPESRAHNTTNETYTVGWSPIYR 197
Db      216  WRCLSKAYLRSGKVVAFDRDDDPWAPLKPALSGVGRWHTTDDVYALGSAGLYR 275

QY      198  TGTSTVNCIVEEDARSVPYSYFAMANGDIANSPEYGLSPPEAAABPMGPQDNFKQLD 257

```

```

Db      276  TGTSTVNCIVEEVARSVTPYDSFALSTGDILYMSPFYGLR-EGAHREHTSYSPERFOQIE 334
QY      258  SYFSMDLDRKRKASLPVKRNFELITSHFTVGVGDWAPKTRVCSMTKWKVETMLKRVNGR 317
Db      335  GYKREDMATGRRLKEPVSRLRTOHVTVANDWPVKRKNVCSLAKWRADEMLRDESRGN 394
QY      318  YRFWARELSATFISNTTEFDNRIILGOCIKREAPAAIEQIFRTKYNDSHVKVGHVQYFL 377
Db      395  FRFTARSLSATFVSDSHTFALQNVPLSDCVITEEAEEAAVERVRYRYNGTHVLSGLSEIYL 454
QY      378  ALGFTIVAYQVPLSKSLAHMYLRELNRNRRDEMLDLVNNKHAIYKKKATSLSRURDRI 437
Db      455  ARGGFVAFRPMLSNELAKLYQLQELARSNGTLEGLFAA---AAPKPGP---RRARRAAP 507
QY      438  NAPNRKITL-----DDTTAITSSTSSVQFAMLOFLYDHIQTHINDMFSTRATACELQN 490
Db      508  SAPGGPGAANGPAGDGDAGGRVTTVSSAEPAALQFTYDHIQDQVHTMFSTRATSCLLQN 567
QY      491  RELVLWHEGIKINPSATASATLGRVAAKMLGDVAASVSSCTAIDAESVTLQNSMRVITST 550
Db      568  KERALWAEAAKLNPSAAASALDRRAAARMGLDAMAVTYCHELGEGRVFIENSMT-APG 625
QY      551  NTCYSRPLVLFSGYENOGNIOGLENELLPTLEAVEPCSANHRRYFLFGSGYALFENY 610
Db      626  GVCISRPVPSFAGNESEPFVGGEDNELLPGRVELPECTANHKRYFRFGADYVYENY 685
QY      611  NFVKMVDAAIDQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVDYAEAVARRNQLHE 670
Db      686  AYRVRVPLAELEVIESTFVDLNLTVLEDREFLPLEVYTRAEALDGLDLYSEIQRRNQLHE 745
QY      671  LKFYDINKVIEVDNYA---GLQEF 692
Db      746  LRFYDIDRVRVKTGNNMIMGLANF 770

RESULT 13
D48349
glycoprotein B precursor - salmeline herpesvirus 1 (strain MV-5-4-PSL)
C:Species: salmeline herpesvirus 1
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A:Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhin
A:Reference number: A48349; MUID:93228440; PMID:8385913
A:Accession: D48349
A:Molecule type: DNA
A:Residues: 1-920 <EBE>
A:Note: sequence extracted from NCBI backbone (NCBIN:1290666, NCBIIP:1290668)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-920/Product: glycoprotein B #status predicted <MAT>
F:576-592/Product: glycoprotein B #status predicted <MAT>
F:576-592/Region: hydrophobic
F:761-780/Domain: transmembrane #status predicted <TM1>
F:787-804/Domain: transmembrane #status predicted <TM2>
F:98,119,152,409,441,683,733/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match      30.7%; Score 1727.5; DB 1; Length 920;
Best Local Similarity 49.1%; Pred. No. 4.8e-79;
Matches 332; Conservative 97; Mismatches 218; Indels 29; Gaps 4;

QY      20  NSSPSTQNTSREVSVSSVQLSEESTFYLCPPPGVSTVIRLEPPRPKCPPEPRKATEWGE 79
Db      98  NKTAEDARAQLROSVRQIRAEANATSMFYVCPPTGATVVOFEPRPCDVAAGKNFTEGI 157

QY      80  AILPKENISPYKFKVLYKNIQTWTGTYTQIRYDTRTPVSEETDLDLIGK 139
Db      158  AVIFKENIAPYKKTATMYKEITVTQWGSRYLQLTGLYNDRAVPFPEITDINAKGL 217

QY      140  CSSKARYLRNNVVEAFDRDAGEKOVLLKPSKENTPESRAHNTTNETYTVGWSPIYRTG 199

```


[illegible]

Search completed: October 8, 2003, 17:08:28
Job time : 41.2016 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:50 ; Search time 20.5086 Seconds
(without alignments)
1552.587 Million cell updates/sec

Title: US-09-147-052-4

Perfect score: 5619

Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNNNAKIPGYRRPGTFL 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 137520 seqs, 29319821 residues

Total number of hits satisfying chosen parameters: 137520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1736	30.9	904	1	PCT-US03-11231-18
2	681	12.1	854	5	US-09-350-841A-1589
3	181.5	3.2	770	5	US-09-897-516A-4453
4	156	2.8	807	5	US-09-820-843B-108
5	151.5	2.7	2663	7	US-60-490-890-1432
6	149	2.7	3418	1	PCT-US03-24394-1
7	147	2.6	1261	1	PCT-US03-26491-192
8	147	2.6	1261	6	US-10-648-593-192
9	146.5	2.6	523	6	US-10-425-114A-58966
10	146.5	2.6	1033	5	US-09-820-843B-75
11	145	2.6	1323	1	PCT-US03-19153-312
12	143.5	2.6	1010	6	US-10-415-657-10
13	141.5	2.5	1979	7	US-60-490-890-1637
14	141	2.5	1427	7	US-60-490-890-824
15	140.5	2.5	386	1	PCT-US03-20322-229
16	140.5	2.5	1454	6	US-10-646-874-22
17	138.5	2.5	1225	5	US-09-581-286A-485
18	138.5	2.5	1226	5	US-09-581-286A-484
19	138.5	2.5	1251	5	US-09-581-286A-351
20	138	2.5	1139	5	US-09-820-843B-15
21	138	2.5	1274	5	US-09-581-286A-466
22	138	2.5	1282	5	US-09-581-286A-465
23	138	2.5	1312	5	US-09-581-286A-338
24	136.5	2.4	1386	6	US-10-472-078-2
25	136.5	2.4	1449	6	US-10-472-078-2
26	136	2.4	402	1	PCT-US02-34769-6

ALIGNMENTS

RESULT 1

PCT-US03-11231-18

; Sequence 18, Application PC/TUS0311231

; GENERAL INFORMATION:

; APPLICANT: Corixa Corporation

; APPLICANT: Day, Craig H.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Parsons, Joseph M.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION

; FILE REFERENCE: 210121.53801PC

; CURRENT APPLICATION NUMBER: PCT/US03/11231

; CURRENT FILING DATE: 2003-04-09

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 18

; LENGTH: 904

; TYPE: PRT

; ORGANISM: Herpes simplex virus

PCT-US03-11231-18

Query Match 30.9%; Score 1736; DB 1; Length 904;
Best Local Similarity 43.0%; Pred. No. 1.8e-79;
Matches 363; Conservative 142; Mismatches 269; Indels 70; Gaps 16;

QY 20 NSPSPQNTVSREVSVSSVOLSEESTFYLCPPPVGVSTVINLEPPKCPPEPRKATEWGEI 79

Db 82 NATVAAGHATLRAHLREIKVENADAQFYVCPPTGATVVFQEQPRRCPTRPEGQNTTEGI 141

QY 80 AILFKENISPYKFKVLYYKNIIQTWTGTTTQITNRYTDRTPTYSIEITLIDGKR 139

Db 142 AVFKENIAPYKATMYIKDVTVSQVWFGRHSQFMGIFEDRAPVFEEVIDKRTKV 201

QY 140 CSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAWHTTNTETVTWVGSPWYRTG 199

Db 202 CRSTAKYVNNMETTAHRRDDHETDMELKPAKATVATSRGWHTTDLKNPSRVEAFHRYG 261

QY 200 TSVNCIVEEMDARSVPYFYFAMANGIANISPFYGLSPPEAAEPMGYPQDNFKOLDYS 259

Db 262 TTVNCIVEEDARSVPYDFEVLATGDFVYMSPFYGR-EGSHTETSVAADRKFQVDGF 320

QY 260 FSDMLDKRKASLPVKNELIISHFTVGVWDNAPKTRVCSWTKKWEVTEMLRATVNGRR 319

Db 321 YARDLTAKARASPTTRNLLTTPKFVAVDWDVPRPAVCTMTKWQVEVDEMLRAEYGSFR 380

QY 320 FWARELSATFISNTTEFDNRIILGOCIKREAAAEQIFRTKYNDSHVKVGHVQVFLAL 379

Db 381 FSSDAISTFTTNLTNLTYSLSRVDLDCIGDRAREADRMFARKYNATHIKVGOPQYLLAT 440

Sequence 14, Appl
Sequence 173, App
Sequence 142, App
Sequence 109, App
Sequence 4, Appl
Sequence 8, Appl
Sequence 1613, Ap
Sequence 26, Appl
Sequence 7, Appl
Sequence 29, Appl
Sequence 25, Appl
Sequence 19, Appl
Sequence 27, Appl
Sequence 23, Appl
Sequence 2, Appl
Sequence 40, Appl
Sequence 38, Appl
Sequence 4698, Ap

QY 380 GGFIYAVQPVLSKSLAHMYLRLNDRNDEMIDLVNKKHAIYKKNATSLRLRRDIRNA 439
 Db 441 GGFIYAVQPVLSKSLAHMYLRLNDRNDEMIDLVNKKHAIYKKNATSLRLRRDIRNA 482
 QY 440 PNRITLDDTTAISTKSTSVFAMQLQYDHIQTHINDMFSRIATAWCELQNLRELVLWHEG 499
 Db 483 PSANASVE--RIKTTSSIFARLOFYNIHQIRVNDMLGRIAVAWCELQNLHELTLWNEA 539
 QY 500 IKINPSATASATLGRRAAKMLGDAVAVSSCTAIDAEVTLQNSMRVITSTNTCYSRPLV 559
 Db 540 RKLNPATASATVGRVRSARMIGDVAVSTCPVAPDNVIVQNSMRVSSRPCTCYSRPLV 599
 QY 560 LFSYGENQGN--IQQLGENNELPTLEAVEPCSNHRRYFLFGSGYALFENYFVKAVDA 618
 Db 600 SFYR--EDQPLIEQOLGENNELRLTRDALEPCTVGHRRYFIFGGYVYFEEYAYSHQLSR 658
 QY 619 ADIQTASTFVELNLTLEDRILPLSVYTKBELRDVGVDYAEVARNQHLKELFYDINK 678
 Db 659 ADVTVSTFDLNTLMLEDHEFVPLEVYTRHEIKDSGLLDYTEVQRNQLHDLRFADIDT 718
 QY 679 VIEVDTN--YAGLOEF-----GCMSTTKDANPNNGQTLQLEAARMELTDLI 722
 Db 719 VIRADANAAMPAGLCAFEFGMDLGRVGVVGVVGVVSAVSGVSSF-----MSNPF 772
 QY 723 NAKAWTLASLDYAKIEASLSAYSEAEFTV---NNLNA---TLQOLKMAKTNLESAL 774
 Db 773 GALAVGLLVL-----AGLVAAPFAFYVQLQORNPMAKALYPLTTKELTSDPGGVGGE 825
 QY 775 NOANTDKTTFDNEHPNLVEA-----YKALKTTLE--QRATNLEGSLSTAYNQIRNLLVDL 827
 Db 826 GEEGAEGGFD--EAKLAAREMIRYALVSAEMTERHKARKGTSALLSSKVTNMVLRK 883
 QY 828 YNKA 831
 Db 884 RNKA 887

RESULT 2

US-09-350-841A-1589
 ; Sequence 1589, Application US/09350841A
 ; GENERAL INFORMATION:
 ; APPLICANT: Jeffs, Peter;
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
 ; FILE REFERENCE: 7872-066-999
 ; CURRENT APPLICATION NUMBER: US/09/350.841A
 ; NUMBER OF SEQ ID NOS: 1999-07-09
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1589
 ; LENGTH: 854
 ; TYPE: PRT
 ; ORGANISM: Human Immunodeficiency virus type 1
 US-09-350-841A-1589

Query Match 12.1%; Score 681; DB 5; Length 854;
 Best Local Similarity 27.7%; Pred. No. 4.7e-27;
 Matches 211; Conservative 122; Mismatches 338; Indels 90; Gaps 21;
 QY 6 RNCIFFLVILYG-----TNSPSTQNTVTSREVSVSVQSEESTFYLCPPVGVSTV 57
 Db 3 RRVLSVVVLLAALACRGAQTPEQPAPPATVQPTATROQTS---PPFRVCELSSHGD 59
 QY 58 IRLEPPRCKPRKATENGEGIALFENISPYKFKVPLYYKNIQTTTWTGTYRQITN 117
 Db 60 FRFSSDIQCSFGRENTEGLLMVFKDNIIPYFVKVRSYTKIVTNILYNGWYADSVTN 119
 QY 118 RYTORPVSIEITDIDLGKRCCKARYLN---NVYVEAFDRDAGEKQVLLKPSKENT 174
 Db 120 RHEKFSVDSTE--TDQMTIYQCNNAVAKMTDGGTTRVTV---DRDGVNITVNLKAPTGG 175
 QY 175 PESRAHWTTNITYTWG--SPWYRTGTSVNCIVBEMDARSFVPSYFAMANGDIANTISPF 233

Db 176 NGVRRYASQTELYDAPGWLITWYTRTIVNCLITDMAKSNSPDFEVTTTGQTVEMSPF 235
 QY 234 YGLSPPEAAAPMGYPQDNFKQLDSYFMDLRRKRKASLPVKRNLFTSHFTVGVWDWAPK 293
 Db 236 YDGKKNKTFPE---RADSPHVRTNYKIVDYNKGTNPQGERAFLDKGYTILSWKLENR 291
 QY 294 TTRVCSMTKKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILG-----QC 346
 Db 292 TA--YCPQLHQWTFDSTIATETGKSIHFVTDGTSSEFTNTT-----VGIELPDAFKC 342
 QY 347 IKREAAAIQIIFRTKYNDSHVKVGHVQ-----YFLALGGFIVAYQPVLSKSLA 395
 Db 343 IE-----EQV-----NKTHEYEAQVDRTYTKGQEAITYITSGGULLLAWLPDTPRSLA 390
 QY 396 HM--YLRELMNRDNRDMLDLVNNKHAITYKKNATSLSLRRDIIRNAPNRKITLDDTTA 454
 Db 391 TVKNLTTELTTTSSPPSSPPAPSA--ARGSTPAAVLRRRRRDAGNATTPVPPTPACKS 448
 QY 455 TSSVQ---FAMQLQYDHIQTHINDMFSRIATAWCELQNLRELVLWHEGKIINPSATASAT 511
 Db 449 LGTLNNPATVVOIOFAYDSLRRQINRMILGDLARAWCLEQKRONMVLRELTKINPTVMSSI 508
 QY 512 LGRVAAAKMGDVAASCTAIDAEVTLQNSMRVITSTNTCYSRPLVLSYSGENQGNQ 571
 Db 509 YGRAVAAKRLGDIYSVQCPVN--QATVLRKMRVPGSETMCSYRPLVSFINDTITYE 567
 QY 572 GQJGENNELPTLEAVEPCSNHRRYFLFGSGYALFENYFVKVMDAADTQIASTFFVELN 631
 Db 568 GOLGTDNEIFLTKMTEVCQATSOYFQSGNEIHVNDYHFKTEILDGIATLQTFIS--N 626
 QY 632 LTLEDREILPLSVYTKBELRDVGVDYAEVARNQHLKELFYDINKVIEVDVNYAGLOE 691
 Db 627 TSLIENIDFASLELYSRDEQRASNVFDEGIFR-----EYNFQ-----AQNIAGLRK 673
 QY 692 FGCMSTTKDANPNNGQTLQLEAARMELTDLINAKAMTSLASL 732
 Db 674 -----DLDNAVSNGRNQFVDGLGELMDSLGVSQISITNL 707

RESULT 3

US-09-897-516A-4453
 ; Sequence 4453, Application US/09897516A
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Huesing, Joseph E.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spiridonov, Sergei
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)B
 ; CURRENT APPLICATION NUMBER: US/09/897.516A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/215,161
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 8415
 ; SEQ ID NO 4453
 ; LENGTH: 770
 ; TYPE: PRT
 ; ORGANISM: Xenorhabdus sp.
 US-09-897-516A-4453

Query Match 3.2%; Score 181.5; DB 5; Length 770;
 Best Local Similarity 20.3%; Pred. No. 0.029;
 Matches 180; Conservative 141; Mismatches 342; Indels 223; Gaps 45;
 QY 257 DSYFSM--DLDKRKASLPVKRNLFTSHFTVGVWDWAPKTRVCSMTKKEVTEMLRATV 314
 Db 3 NKYFALITDSCANKLENAATSGNKLEITHMAVGDLPTPDANOSQ-----LI 50
 QY 315 NGVRRFMARELSATFISNTTEFDPNRIILGQCI--KREAAAIQIIFRTKYNDSHVKVGH- 372

Qy	488	LONRELVWHEG	KINFSATASATLGRVAAKMLGDVA	SSCTDAESTVLQNSMRVI	547
Db	1486	LKEOETINELRVNL	SEKETEISTIQKQ	-----LEAINDKLQNKIOEI	1528
Qy	548	TSTNTCYSRPLVLF	SYCENOGNIQOG	GENNELPTLEAVEPCSANHRRYFLPGSGYALF	607
Db	1529	-----YEK	-----EEOGLNK	QISEVOENNELKQPK-----EHR	1557
Qy	608	ENYFVKWVDAADI	QIATSTFVEKLNLTLEDREILPLSVYTK	EELDVGVLDYAEVARRNQ	667
Db	1558	-----KAKDSALQ	SIESKMLELTNRLQESQEBIQIMIKEKEMKV	-----QEAQLQIERDQ	1608
Qy	668	LHE	-----LKFDYINKVIEVDNYAGLQEGFGCS	ITKKDANPNNGQ	708
Db	1609	LKENTREIVAKM	KEQKEYQFLKMTAVNETQKMCIEHLKE	-----QFETQKLNLEN	1662
Qy	709	TOLEAARMELTD	LNAKAMTLASI	-----QDYAKIEASLSAYSEATVNNLNATC	760
Db	1663	--IETENIRLTQIL	HENLEEMRSVTKERDRLSRVEETLK	-----VEROQLKENLRETTYRDL	1717
Qy	761	---EOLKMAKTNL	ESAINQANTK	-----TTFDNEHNPVLVEAYKALKTTL	802
Db	1718	EKOELKIVHMLKE	--HQETIDKLRGIVSEKTEI	SNMQDLHSN--DALKAQDLKIQ	1773
Qy	803	-----EQRAT	--NLEGSSATYNOIRNNLV	DLYNKASSLITKTLDPPLNGTILD	849
Db	1774	EELRIAHHMLKE	OQETIDKLRGIVSEKTKLSNMQDL	ENSNAKLOEKIOE-----LK	1826
Qy	850	SNE--TWTANKN	NNLTSTINEOKTNADAL	SNFSIKKVION	888
Db	1827	ANEHQITLTKD	YNETQKKVSEMEQ	-----LKKOIKD	1858

```

RESULT 6
PCT-US03-24394-1
; Sequence 1, Application PC/TUS0324394
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute
; APPLICANT: Ramin Shiekhattar
; TITLE OF INVENTION: METHODS FOR REGULATING BRCA1-BRCA2-CONTAINING COMPLEX ACTIVITY
; FILE REFERENCE: WSTR-0014B
; CURRENT APPLICATION NUMBER: PCT/US03/24394
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/401,433
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/449,950
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3418
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-24394-1

```

Query Match	2.7%;	Score 149;	DB 1;	Length 3418;
Best Local Similarity	18.1%;	Pred. No. 5.3;		
Matches	213;	Conservative 189;	Mismatches 396;	Indels 378; Gaps 59;
QY	85	ENISPK--FKVLYLYKNIQTWTGTYTQRTNRYTDRTPVSIEITLDIDGRCSS	142	
Db	1625	ENLTKSKSIFLKVKVHENVEKETAKSPATC-----YTNSPYSVIENSALAFVTS--CSR	1677	
QY	143	KA-----RYLRNNVYVEAFORDAGEQVLLKPKSFNPESRAWHTTNYTWG	191	
Db	1678	KTSVSOTSLEAKKWLRGII-----PG-----OPERINAD-----YV	1711	
QY	192	SPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANTSPFYGLSPPEAAAPMGYPQD	251	
Db	1712	GNLYENNS--NSTIAENDKNHLSEKQDIYLSNSSMSNSYSYHS-----DEVYNDSGYLSK	1765	
QY	252	NFKQLDYSYSDML-----DKRRKASLPVKKNFLITSHFTVGWDWAPKTR--VCSMTKKKE	305	

[illegible]

RESULT 7
PCT-US03-26491-192
; Sequence 192, Application PC/TUS0326491
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS

; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/26491
; CURRENT FILING DATE: 2003-08-26
; PRIOR FILING DATE: 2003-08-26
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-26491-192

Query Match 2.6%; Score 147; DB 1; Length 1261;
Best Local Similarity 19.3%; Pred. No. 2.5;
Matches 136; Conservative 109; Mismatches 254; Indels 206; Gaps 34;

QY 307 TEMPLRATVNGRYR--FMARELSATFISNTTEPDNRIILGQCICKREAAAEIQIFRTKYN 364
Db 9 TKKRAWASQGLSTDTTSEMGLKSLSSNSIFDPDYI-----KELVN 50
QY 365 DSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDRNTDEMLDVNNKHAIYKK 424
Db 51 DIR-KFSHILLYLKEAIFSDCFKEVI-----HIRLEELLR-----VLKSIIMKH-----Q 94
QY 425 NATSLRLRRDIRNAPNRKITDDTTAISTSSVQFAMQLFYDHIQTHINDMFSRIATA 484
Db 95 NLNSV-----DLQNAEML-----TAKVK---AVNFEVN-----EENKNDLFQEV--- 132
QY 485 WCBLQNLRELVLWHEGKIINPSATATLGRVAAKMLGDVAAYSSC-TAIDAESVTQLN- 542
Db 133 -----FSSIEFLAFTFGNLTFLMGDVGNDSEFLRLPVSRETKEFENV 175
QY 543 SMRVITSTNTCYSRPLVLSYGENOQIQLGNNELLPTLEAVEPCSANHRRYFLFGS 602
Db 176 SVESVDSSS-----EKGNF-SPLELDNVLLKNTDSIE-----LAL 209
QY 603 GYA-LFENY--NFKVMDAADIQIASTFVELNLTLEDRILPLSVYTTKEELRDVGVDY 659
Db 210 SYAKTWSKYTKNIVSVWEK-----KLNLELESTRNMVKLAETR---TNIGIQEF 256
QY 660 AEVARNQHLKIFYDINKVIEVDVTNYAGLQEGFCHMSITKKDANPNNG-OTQLEAARMEL 718
Db 257 MPL--OSLFTNALLNDIESHLLQQTIAALQ-----ANKFVQPLLGRKNEMEKORKEI 307
QY 719 TD-----LINAKAMTLASLDQYAKIEASLSSA-----YSEATVNNNLNAT 759
Db 308 KELWKQEQNMKLEAEALAKKALLCMORODEYEKAKSSMFRAEEHLSGGGLAKNLNQ 367
QY 760 LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQ 819
Db 368 LEKKRLEEEALQKVEAD-----ELYKVCVTNVEERNDVENTKREILAQ 413
QY 820 IRNNL--VDLYNKASSLITKTLDPNGTLLDSNEITTANKNINNTLSTINEQKTADAL 877
Db 414 LRTLVPQCOLDTLKAVTNVLFHMOHLQAASLADRLQSLGSAKLYDPQGEYSE----- 465
QY 878 SNSFIKKVIONNEQSFVGFTTNAN-----VQPSNYSFVAESADVTVPNTKYARRTVWNGD 932
Db 466 ---FVKATNSTEEKVDG---NVNKHNLSSQPSGFG-----PAN---SLEDVVRLP 507
QY 933 EPSSRI-----LANTNSITDVSWIYSIAGTNTKYQFSFSGVSTG 973
Db 508 DSSNKIEEDRCNSADITGFSFIRS-----WTFGMFSDSESTG 545

RESULT 8

US-10-648-593-192
; Sequence 192, Application US/10648593
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company

RESULT 9

US-10-425-114A-58966
; Sequence 58966, Application US/10425114A
; GENERAL INFORMATION:

; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS T
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-192

Query Match 2.6%; Score 147; DB 6; Length 1261;
Best Local Similarity 19.3%; Pred. No. 2.5;
Matches 136; Conservative 109; Mismatches 254; Indels 206; Gaps 34;

QY 307 TEMPLRATVNGRYR--FMARELSATFISNTTEPDNRIILGQCICKREAAAEIQIFRTKYN 364
Db 9 TKKRAWASQGLSTDTTSEMGLKSLSSNSIFDPDYI-----KELVN 50
QY 365 DSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMDRNTDEMLDVNNKHAIYKK 424
Db 51 DIR-KFSHILLYLKEAIFSDCFKEVI-----HIRLEELLR-----VLKSIIMKH-----Q 94
QY 425 NATSLRLRRDIRNAPNRKITDDTTAISTSSVQFAMQLFYDHIQTHINDMFSRIATA 484
Db 95 NLNSV-----DLQNAEML-----TAKVK---AVNFEVN-----EENKNDLFQEV--- 132
QY 485 WCBLQNLRELVLWHEGKIINPSATATLGRVAAKMLGDVAAYSSC-TAIDAESVTQLN- 542
Db 133 -----FSSIEFLAFTFGNLTFLMGDVGNDSEFLRLPVSRETKEFENV 175
QY 543 SMRVITSTNTCYSRPLVLSYGENOQIQLGNNELLPTLEAVEPCSANHRRYFLFGS 602
Db 176 SVESVDSSS-----EKGNF-SPLELDNVLLKNTDSIE-----LAL 209
QY 603 GYA-LFENY--NFKVMDAADIQIASTFVELNLTLEDRILPLSVYTTKEELRDVGVDY 659
Db 210 SYAKTWSKYTKNIVSVWEK-----KLNLELESTRNMVKLAETR---TNIGIQEF 256
QY 660 AEVARNQHLKIFYDINKVIEVDVTNYAGLQEGFCHMSITKKDANPNNG-OTQLEAARMEL 718
Db 257 MPL--OSLFTNALLNDIESHLLQQTIAALQ-----ANKFVQPLLGRKNEMEKORKEI 307
QY 719 TD-----LINAKAMTLASLDQYAKIEASLSSA-----YSEATVNNNLNAT 759
Db 308 KELWKQEQNMKLEAEALAKKALLCMORODEYEKAKSSMFRAEEHLSGGGLAKNLNQ 367
QY 760 LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQ 819
Db 368 LEKKRLEEEALQKVEAD-----ELYKVCVTNVEERNDVENTKREILAQ 413
QY 820 IRNNL--VDLYNKASSLITKTLDPNGTLLDSNEITTANKNINNTLSTINEQKTADAL 877
Db 414 LRTLVPQCOLDTLKAVTNVLFHMOHLQAASLADRLQSLGSAKLYDPQGEYSE----- 465
QY 878 SNSFIKKVIONNEQSFVGFTTNAN-----VQPSNYSFVAESADVTVPNTKYARRTVWNGD 932
Db 466 ---FVKATNSTEEKVDG---NVNKHNLSSQPSGFG-----PAN---SLEDVVRLP 507
QY 933 EPSSRI-----LANTNSITDVSWIYSIAGTNTKYQFSFSGVSTG 973
Db 508 DSSNKIEEDRCNSADITGFSFIRS-----WTFGMFSDSESTG 545

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:00 ; Search time 434.972 Seconds
(without alignments)
2271.804 Million cell updates/sec

Title: US-09-147-052-4

Perfect score: 5619

Sequence: 1 MHYFRRNCIFLIIVLGTN.....SSNENADKIPYRRPGTFL 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	5619	100.0	1086	15 US-09-147-052-4

Sequence 4, Appl

2	3601	64.1	865	3	US-07-722-860-13	Sequence 13, Appl
3	3601	64.1	865	6	US-08-213-449A-13	Sequence 13, Appl
4	3601	64.1	865	6	US-08-213-449B-13	Sequence 13, Appl
5	2015	35.9	456	15	US-09-147-052-2	Sequence 2, Appl
6	2008	35.7	456	24	US-09-901-572A-3	Sequence 3, Appl
7	1929.5	34.3	384	27	US-10-131-591A-12	Sequence 12, Appl
8	1777	31.6	357	27	US-10-131-591A-23	Sequence 23, Appl
9	1775.5	31.6	868	12	US-08-804-439-21	Sequence 21, Appl
10	1775.5	31.6	868	17	US-09-301-390-21	Sequence 21, Appl
11	1775.5	31.6	868	17	US-09-338-326-21	Sequence 21, Appl
12	1749	31.1	933	3	US-07-805-524-2	Sequence 2, Appl
13	1743	31.0	943	4	US-08-096-183D-4	Sequence 4, Appl
14	1743	31.0	943	19	US-09-521-738-14	Sequence 14, Appl
15	1739.5	31.0	891	9	US-08-541-878-6	Sequence 6, Appl
16	1736	30.9	904	1	PCT-US03-11231-18	Sequence 18, Appl
17	1736	30.9	904	27	US-10-121-988-18	Sequence 18, Appl
18	1736	30.9	904	28	US-10-200-562-18	Sequence 18, Appl
19	1736	30.9	904	28	US-10-237-551-18	Sequence 18, Appl
20	1719.5	30.6	885	9	US-08-541-878-8	Sequence 8, Appl
21	1714.5	30.5	885	12	US-08-804-439-23	Sequence 23, Appl
22	1714.5	30.5	885	17	US-09-301-390-23	Sequence 23, Appl
23	1714.5	30.5	885	17	US-09-338-326-23	Sequence 23, Appl
24	1702.5	30.3	903	12	US-08-804-439-22	Sequence 22, Appl
25	1702.5	30.3	903	17	US-09-301-390-22	Sequence 22, Appl
26	1702.5	30.3	903	17	US-09-338-326-22	Sequence 22, Appl
27	1696.5	30.2	904	32	US-60-412-956-12	Sequence 12, Appl
28	1695	30.2	846	5	US-08-123-456-142	Sequence 142, App
29	1695	30.2	846	16	US-09-297-477A-142	Sequence 142, App
30	1695	30.2	846	25	US-09-994-404-142	Sequence 142, App
31	1695	30.2	896	5	US-08-123-456-241	Sequence 241, App
32	1695	30.2	896	16	US-09-297-477A-241	Sequence 241, App
33	1695	30.2	896	25	US-09-994-404-241	Sequence 241, App
34	1664.5	29.6	787	16	US-08-123-456-79	Sequence 79, Appl
35	1664.5	29.6	787	16	US-09-297-477A-79	Sequence 79, Appl
36	1664.5	29.6	787	25	US-09-994-404-79	Sequence 79, Appl
37	778.5	13.9	907	4	US-08-048-978-2	Sequence 2, Appl
38	778.5	13.9	907	12	US-08-804-439-19	Sequence 19, Appl
39	778.5	13.9	907	17	US-09-301-390-19	Sequence 19, Appl
40	778.5	13.9	907	17	US-09-338-326-19	Sequence 19, Appl
41	778.5	13.9	907	28	US-10-223-538-2	Sequence 2, Appl
42	758.5	13.5	830	26	US-10-055-364-45	Sequence 45, Appl
43	756	13.5	906	32	US-60-435-549-11	Sequence 11, Appl
44	754	13.4	865	1	PCT-US00-18647-24	Sequence 24, Appl
45	754	13.4	865	26	US-10-055-364-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-147-052-4

; Sequence 4, Application US/09147052

; GENERAL INFORMATION:

; APPLICANT: SAITOH, Shuji

; APPLICANT: TSUZAKI, Yoshinari

; APPLICANT: YANAGIDA, Noboru

; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR, RECOMBINANT VIRUS, AND ITS USE

; FILE REFERENCE: 981167

; CURRENT APPLICATION NUMBER: US/09/147,052

; CURRENT FILING DATE: 1999-04-05

; PRIOR APPLICATION NUMBER: JP 08-103548

; PRIOR FILING DATE: 1996-03-29

; PRIOR APPLICATION NUMBER: PCT/JP97/01084

; PRIOR FILING DATE: 1997-03-28

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn ver. 2.1

; SEQ ID NO 4

; LENGTH: 1086

; TYPE: PRT

; ORGANISM: hybrid

US-09-147-052-4

Query Match		100.0%;	Score 5619;	DB 15;	Length 1086;				
Best Local Similarity		100.0%;	Pred. No. 0;						
Matches 1086;		Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MHYFRRNCIFFLIVILYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGSTVIRL	60						
Db	1	MHYFRRNCIFFLIVILYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGSTVIRL	60						
QY	61	EPKRCPEPRKATWEGGIALFKENISPYKFKVTLTKYKNIQTWTGTYRQITNRYT	120						
Db	61	EPKRCPEPRKATWEGGIALFKENISPYKFKVTLTKYKNIQTWTGTYRQITNRYT	120						
QY	121	DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW	180						
Db	121	DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW	180						
QY	181	HTTNETYTVWGSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPFYGLSPPE	240						
Db	181	HTTNETYTVWGSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPFYGLSPPE	240						
QY	241	AAAEPGMPQDNFKQDYSFMSDLDRKRSALPVRKRNFLITSHTVGDWNAKTRVCSM	300						
Db	241	AAAEPGMPQDNFKQDYSFMSDLDRKRSALPVRKRNFLITSHTVGDWNAKTRVCSM	300						
QY	301	TKWKEVTEMLRATVNGRYRFWARELSATFISNTTEFPDNRILIGQCIRKREAAIEQIFR	360						
Db	301	TKWKEVTEMLRATVNGRYRFWARELSATFISNTTEFPDNRILIGQCIRKREAAIEQIFR	360						
QY	361	TKYNDSHVKGHVQYFLALGFIIVAYQPVLSKSLAHMYLRMLRNDRTDEMLDVNNKHA	420						
Db	361	TKYNDSHVKGHVQYFLALGFIIVAYQPVLSKSLAHMYLRMLRNDRTDEMLDVNNKHA	420						
QY	421	IYKKNATSLSLRRDIRNAPNRKLTDDTTAIKTSSTSVQFAMQLQFLYDHIQTHINDMFSR	480						
Db	421	IYKKNATSLSLRRDIRNAPNRKLTDDTTAIKTSSTSVQFAMQLQFLYDHIQTHINDMFSR	480						
QY	481	IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL	540						
Db	481	IATAWCELQNLRELVLWHEGINKINPSATASATLGRVAAKMLGDVAAVSSCTAIDAESVTL	540						
QY	541	QNSMRVITSTNCTSRPLVLFSYGENOGNTOGQGENNELLPTEAVEPCSANHRRYFLF	600						
Db	541	QNSMRVITSTNCTSRPLVLFSYGENOGNTOGQGENNELLPTEAVEPCSANHRRYFLF	600						
QY	601	GSYALFENYFKVMDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA	660						
Db	601	GSYALFENYFKVMDAADIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA	660						
QY	661	EVARRNOLHELKFDYDINKVIEVDITNYAGLQEFQCMSTTKKDANPNNGQTOLEAARMELTD	720						
Db	661	EVARRNOLHELKFDYDINKVIEVDITNYAGLQEFQCMSTTKKDANPNNGQTOLEAARMELTD	720						
QY	721	LINAKAMTLASLDQYAKIEASLSAYSEAEVTANNLNATLEQLKMAKTNLESAINQANTD	780						
Db	721	LINAKAMTLASLDQYAKIEASLSAYSEAEVTANNLNATLEQLKMAKTNLESAINQANTD	780						
QY	781	KTFDNEHPNLVAYKALKTTLEORATNLEGLSTAYNQIRNRLVDLYNKASSLITKTLTD	840						
Db	781	KTFDNEHPNLVAYKALKTTLEORATNLEGLSTAYNQIRNRLVDLYNKASSLITKTLTD	840						
QY	841	PLNGGTLDSNEITATANKNINNTLSTINEQKTNADALSNSFIKKVIQNNQSFVGTFTNA	900						
Db	841	PLNGGTLDSNEITATANKNINNTLSTINEQKTNADALSNSFIKKVIQNNQSFVGTFTNA	900						
QY	901	NVQPSNYSFAVSADVTPVNYKARTVWNGDEPSSRLANTNSITDVSIIYSLAGNTK	960						
Db	901	NVQPSNYSFAVSADVTPVNYKARTVWNGDEPSSRLANTNSITDVSIIYSLAGNTK	960						
QY	961	YQSFSGNYGPGSTGLYFPYKLVKAADANNVGLQYKLNNGVNVQVVEATSTSANNTANPT	1020						
Db	961	YQSFSGNYGPGSTGLYFPYKLVKAADANNVGLQYKLNNGVNVQVVEATSTSANNTANPT	1020						
QY	1021	PAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKNVAPMIGNIYLSSENNAADKIPGYR	1080						
Db	1021	PAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKNVAPMIGNIYLSSENNAADKIPGYR	1080						
RESULT 2									
US-07-722-860-13									
; Sequence 13, Application US/07722860									
; GENERAL INFORMATION:									
; APPLICANT: NAZERIAN, Keyvan									
; APPLICANT: LEE, Lucy F.									
; APPLICANT: YANAGIDA, Noboru									
; APPLICANT: OGAWA, Ryohei									
; APPLICANT: LI, Yi									
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR									
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE									
; NUMBER OF SEQUENCES: 19									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH									
; STREET: 301 North Washington Street									
; CITY: Falls Church									
; STATE: Virginia									
; COUNTRY: USA									
; ZIP: 22040-0747									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Floppy disk									
; COMPUTER: IBM PC compatible									
; OPERATING SYSTEM: PC-DOS/MS-DOS									
; SOFTWARE: PatentIn Release #1.0, Version #1.25									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/07722.860									
; FILING DATE: 19910628									
; CLASSIFICATION: 424									
; ATTORNEY/AGENT INFORMATION:									
; NAME: Murphy Jr., Gerald M.									
; REGISTRATION NUMBER: 28,977									
; REFERENCE/DOCKET NUMBER: 1644-101P									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: (703) 241-1300									
; TELEFAX: (703) 241-2848									
; TELEX: 248345									
; INFORMATION FOR SEQ ID NO: 13:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 865 amino acids									
; TYPE: AMINO ACID									
; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
US-07-722-860-13									
Query Match 64.1%; Score 3601; DB 3; Length 865;									
Best Local Similarity 85.9%; Pred. No. 1.7e-302;									
Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;									
QY	1	MHYFRRNCIFFLIVILYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGSTVIRL	60						
Db	1	MHYFRRNCIFFLIVILYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPPVGSTVIRL	60						
QY	61	EPKRCPEPRKATWEGGIALFKENISPYKFKVTLTKYKNIQTWTGTYRQITNRYT	120						
Db	61	EPKRCPEPRKATWEGGIALFKENISPYKFKVTLTKYKNIQTWTGTYRQITNRYT	120						
QY	121	DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW	180						
Db	121	DRTPVSTEEITDLIDGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW	180						
QY	181	HTTNETYTVWGSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPFYGLSPPE	240						
Db	181	HTTNETYTVWGSPIWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPPFYGLSPPE	240						
QY	241	AAAEPGMPQDNFKQDYSFMSDLDRKRSALPVRKRNFLITSHTVGDWNAKTRVCSM	300						
Db	241	AAAEPGMPQDNFKQDYSFMSDLDRKRSALPVRKRNFLITSHTVGDWNAKTRVCSM	300						

```
Db 241 AAAPGTPQDNFKQDYSFMDLDRKRSALPVARNFLLTSHFTVGVMDWAPKTRVCSM 300
QY 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKEAATAEQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKEAATAEQIFR 360
QY 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMIDLNNKHA 420
Db 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMIDLNNKHA 420
QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLDYHIOHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLDYHIOHIQTHINDMFSR 480
QY 481 IATWCELQNLRELVLWHEGKIKINPSATASATLGRVRAAKMLGDDVAAVSSCTAIDAESVTL 540
Db 481 IATWCELQNLRELVLWHEGKIKINPSATASATLGRVRAAKMLGDDVAAVSSCTAIDAESVTL 540
QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKVMDAADIOIASTFVELNLTLLDREILPLSVYTKKEELRDVGVLDA 660
Db 601 GSGYALFENYFVKVMDAADIOIASTFVELNLTLLDREILPLSVYTKKEELRDVGVLDA 660
QY 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
Db 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
QY 717 ELTDLINKAMTASLQDYAK-----IEASLSAYSEATVN-----NNLNA--- 758
Db 717 ELTDLINKAMTASLQDYAK-----IEASLSAYSEATVN-----NNLNA--- 758
QY 759 -TLEOLKMAKTNLESAINQANTDKTTFDNEHNLVEA-----YKALKITLEQATNL 809
Db 759 -TLEOLKMAKTNLESAINQANTDKTTFDNEHNLVEA-----YKALKITLEQATNL 809
QY 769 MTEVLKQAQATRELHGEESDDELTERTSIDER--KLEAREMIKYMALVSAREERHEKL 823
Db 769 MTEVLKQAQATRELHGEESDDELTERTSIDER--KLEAREMIKYMALVSAREERHEKL 823
```

RESULT 3

US-08-213-449A-13

```
; Sequence 13, Application US/08213449A
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohel
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; City: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213,449A
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-108P
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-213-449A-13
```

Query Match 64.1%; Score 3601; DB 6; Length 865;

Best Local Similarity 85.9%; Pred No 1.7e-302;

Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;

```
QY 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVQLSEBESTFYLCPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVQLSEBESTFYLCPPVGVSTVIRL 60
QY 61 EPPRCPCPRKATENGEGIAILFKENISPYKFKVLYYKNIIQTTWTGTTTQITNRYT 120
Db 61 EPPRCPCPRKATENGEGIAILFKENISPYKFKVLYYKNIIQTTWTGTTTQITNRYT 120
QY 121 DRTPVYSIEITLIDGKRCSSKARYLRNNVVEAFDRDAGBKQVLLKPSKENTPESRAW 180
Db 121 DRTPVYSIEITLIDGKRCSSKARYLRNNVVEAFDRDAGBKQVLLKPSKENTPESRAW 180
QY 181 HTTNETYTVWGSPIWYRTGTSYNCIVEEMDARSFPPSYFAMANGDIANISPFYGLSPPE 240
Db 181 HTTNETYTVWGSPIWYRTGTSYNCIVEEMDARSFPPSYFAMANGDIANISPFYGLSPPE 240
QY 241 AAAPGTPQDNFKQDYSFMDLDRKRSALPVARNFLLTSHFTVGVMDWAPKTRVCSM 300
Db 241 AAAPGTPQDNFKQDYSFMDLDRKRSALPVARNFLLTSHFTVGVMDWAPKTRVCSM 300
QY 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKEAATAEQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRFMARELSATFISNTTFFDPNRIILGQCIRKEAATAEQIFR 360
QY 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMIDLNNKHA 420
Db 361 TKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMIDLNNKHA 420
QY 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLDYHIOHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAKTSSTSSVQFAMQLDYHIOHIQTHINDMFSR 480
QY 481 IATWCELQNLRELVLWHEGKIKINPSATASATLGRVRAAKMLGDDVAAVSSCTAIDAESVTL 540
Db 481 IATWCELQNLRELVLWHEGKIKINPSATASATLGRVRAAKMLGDDVAAVSSCTAIDAESVTL 540
QY 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLFYSGENQNIQOLGENNELPTLEAVEPCSANHRRYFLF 600
QY 601 GSGYALFENYFVKVMDAADIOIASTFVELNLTLLDREILPLSVYTKKEELRDVGVLDA 660
Db 601 GSGYALFENYFVKVMDAADIOIASTFVELNLTLLDREILPLSVYTKKEELRDVGVLDA 660
QY 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
Db 661 EVARRNQLHELKFYDINKVIEVDNTYA---GLQFEGCMSITKKDANPNNGOTQL-EAARM 716
QY 717 ELTDLINKAMTASLQDYAK-----IEASLSAYSEATVN-----NNLNA--- 758
Db 717 ELTDLINKAMTASLQDYAK-----IEASLSAYSEATVN-----NNLNA--- 758
QY 759 -TLEOLKMAKTNLESAINQANTDKTTFDNEHNLVEA-----YKALKITLEQATNL 809
Db 759 -TLEOLKMAKTNLESAINQANTDKTTFDNEHNLVEA-----YKALKITLEQATNL 809
QY 769 MTEVLKQAQATRELHGEESDDELTERTSIDER--KLEAREMIKYMALVSAREERHEKL 823
Db 769 MTEVLKQAQATRELHGEESDDELTERTSIDER--KLEAREMIKYMALVSAREERHEKL 823
```

RESULT 4

US-08-213-449B-13

```

; Sequence 13, Application US/08213449B
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/213,449B
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-108P
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-213-449B-13

```

```

Query Match      64.1%; Score 3601; DB 6; Length 865;
Best Local Similarity 85.9%; Pred. No. 1.7e-302;
Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;

Qy 1 MHYFRNCIFFLVILYGTSSPSTQNTSVREVSSVQLSEESTFYLCPPPVGSTVIRL 60
Db 1 MHYFRNCIFFLVILYGTSSPSTQNTSVREVSSVQLSEESTFYLCPPPVGSTVIRL 60

Qy 61 EPRKCEPRKATEWGGIAILFKENISPYKFKVTLVYKNIOTTTWTGTYRQITNRYT 120
Db 61 EPRKCEPRKATEWGGIAILFKENISPYKFKVTLVYKNIOTTTWTGTYRQITNRYT 120

Qy 121 DRTPVSTEEITDLIDGRCSSKARYLRNNVYVEAFDRDAGEQVLLKPSKFNTPE$RAW 180
Db 121 DRTPVSTEEITDLIDGRCSSKARYLRNNVYVEAFDRDAGEQVLLKPSKFNTPE$RAW 180

Qy 181 HTTNETYVWGSPWIRYTGTSVNCIVEEMDARSVPYSYFAMANGDIANTISPYGLSPPE 240
Db 181 HTTNETYVWGSPWIRYTGTSVNCIVEEMDARSVPYSYFAMANGDIANTISPYGLSPPE 240

Qy 241 AAAPMGYPQDNFKQDYSFMDLDRKRSKASLPVKNRFLITSHTFTVGWDWAPKTRVCSM 300
Db 241 AAAPMGYPQDNFKQDYSFMDLDRKRSKASLPVKNRFLITSHTFTVGWDWAPKTRVCSM 300

Qy 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILQCQIKREAAAEIQIFR 360
Db 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTEFDPNRIILQCQIKREAAAEIQIFR 360

Qy 361 TKYND$HVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLVNKHA 420
Db 361 TKYND$HVKGHVQYFLALGGFIVAYQPVLSKSLAHMYLRELNRDNTDEMLDLVNKHA 420

```

```

Qy 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAIKSTSSVQFAMQLQFLYDHIQTHINDMFSR 480
Db 421 IYKKNATSLRLRRDIRNAPNRKITLDDTTAIKSTSSVQFAMQLQFLYDHIQTHINDMFSR 480

Qy 481 IATAWCELQRELVLWHEGIRKINPSATASATLGRVAAKMLGDVAASCTAIDAESVTL 540
Db 481 IATAWCELQRELVLWHEGIRKINPSATASATLGRVAAKMLGDVAASCTAIDAESVTL 540

Qy 541 QNSMRVITSTNTCYSRPLVLSYGENQNIQQLGNNELLPTLEAVEPC$ANHRRYFLF 600
Db 541 QNSMRVITSTNTCYSRPLVLSYGENQNIQQLGNNELLPTLEAVEPC$ANHRRYFLF 600

Qy 601 GSGYALPENYNFVKMVDAAIQIASTFVELNLTLEDEIREILPLSYTTEELRDVGLDYA 660
Db 601 GSGYALPENYNFVKMVDAAIQIASTFVELNLTLEDEIREILPLSYTTEELRDVGLDYA 660

Qy 661 EVARRNQLHELKFDYDINKVIEVDNYA---GLQEFQCM$ITKDKANPNNGOTQL-EAARM 716
Db 661 EVARRNQLHELKFDYDINKVIEVDNYAFPMGLAEL-----FNGHGOVQQAIGK 708

Qy 717 ELTDLINAKAMTASLQDYAK-----IEASLSSAYSEAEVTN---NNLNA--- 758
Db 717 ELTDLINAKAMTASLQDYAK-----IEASLSSAYSEAEVTN---NNLNA--- 758

Qy 759 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQRATNL 809
Db 759 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQRATNL 809

Qy 809 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQRATNL 809
Db 809 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQRATNL 809

```

```

RESULT 5
US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
; US-09-147-052-2

```

```

Query Match      35.9%; Score 2015; DB 15; Length 456;
Best Local Similarity 100.0%; Pred. No. 3.8e-165;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 691 EFGCM$ITKDKANPNNGOTQLEAARMELTDLINAKAMTASLQDYAKIEASLSSAYSEAE 750
Db 61 EFGCM$ITKDKANPNNGOTQLEAARMELTDLINAKAMTASLQDYAKIEASLSSAYSEAE 120

Qy 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 810
Db 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180

Qy 811 GLSSTAYNQNRNLVDLYNKASSLITKTLQPLNGTLLD$NEITANKNNINTLSTINEQ 870
Db 181 GLSSTAYNQNRNLVDLYNKASSLITKTLQPLNGTLLD$NEITANKNNINTLSTINEQ 240

Qy 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAF$ADVTPVNYKYARRVWN 930
Db 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAF$ADVTPVNYKYARRVWN 300

```


QY 931 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 990
DB 301 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 360
QY 991 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
DB 361 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
QY 1051 EGNMNVAPMIGNIYLSNENNAKIPGYRRPGTFL 1086
DB 421 EGNMNVAPMIGNIYLSNENNAKIPGYRRPGTFL 456

RESULT 6
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PR1
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pN240K-S
US-09-901-572A-3

Query Match 35.7%; Score 2008; DB 24; Length 456;
Best Local Similarity 99.5%; Pred. No. 1.5e-164;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 691 EFGCMSITKKDANPNNGQTLAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAE 750
DB 61 EFGCMSITKKDANPNNGQTLAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAE 120
QY 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLE 810
DB 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLE 180
QY 811 GLSSTAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTTANKNINNTLSTINEQ 870
DB 181 GLSSTAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTTANKNINNTLSTINEQ 240
QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWN 930
DB 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWN 300
QY 931 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 990
DB 301 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 360
QY 991 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
DB 361 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
QY 1051 EGNMNVAPMIGNIYLSNENNAKIPGYRRPGTFL 1086
DB 421 EGNMNVAPMIGNIYLSNENNAKIPGYRRPGTFL 456

RESULT 7
US-10-131-591A-12
; Sequence 12, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A

; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PR1
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pN240K-S
US-10-131-591A-12

Query Match 34.3%; Score 1929.5; DB 27; Length 384;
Best Local Similarity 99.2%; Pred. No. 7.5e-158;
Matches 382; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 691 EFGCMSITKKDANPNNGQTLAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAE 750
DB 1 EF-CMSITKKDANPNNGQTLAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAE 59
QY 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLE 810
DB 60 TVNNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLE 119
QY 811 GLSSTAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTTANKNINNTLSTINEQ 870
DB 120 GLSSTAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTTANKNINNTLSTINEQ 179
QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWN 930
DB 180 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWN 239
QY 931 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 990
DB 240 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKAADANNV 299
QY 991 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
DB 300 GLQYKLNNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 359
QY 1051 EGNMNVAPMIGNIYLSNENNAK 1075
DB 360 EGNMNVAPMIGNIYLSNENNAK 384

RESULT 8
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PR1
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of
; OTHER INFORMATION: pN240K-S
US-10-131-591A-23

Query Match 31.6%; Score 1777; DB 27; Length 357;
Best Local Similarity 98.6%; Pred. No. 1.2e-144;
Matches 352; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 719 TDLINAKAMTILASLDQYAKIEASLSAYSEAEETVNNLNATLEQLKMAKTNLESAINQAN 778
DB 1 TDLINAKAMTILASLDQYAKIEASLSAYSEAEETVNNLNATLEQLKMAKTNLESAINQAN 60
QY 779 TDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNOIRNNLVLYNKASSLITKT 838

```
Db      61 TDKTFDEHNLVAYKALKTLEORATNLEGSTAYNQIRNLVLDYKASLLTKT 120
QY      839 LDPLNGGTLDSNEITTTANKNINNTLSTINBQKTNADALSNSFIKKVIONNEQSFVGTF 898
Db      121 LDPLNGGTLDSNEITTTANKNINNTLSTINBQKTNADALSNSFIKKVIONNEQSFVGTF 180
QY      899 NANYQPSNYSFVASADYTPYNYKYARTVWNGDEPSSRIILANTNSITDVSIIYSLAGTN 958
Db      181 NANYQPSNYSFVASADYTPYNYKYARTVWNGDEPSSRIILANTNSITDVSIIYSLAGTN 240
QY      959 TKYQSFSGNYGPGSYLYFPKLYKAADANNVGLQYKLNNGNVQOQVEFATSTSNNTTAN 1018
Db      241 TKYQSFSGNYGPGSYLYFPKLYKAADANNVGLQYKLNNGNVQOQVEFATSTSAQQTAN 300
QY      1019 PTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADK 1075
Db      301 PTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADK 357

RESULT 9
US-08-804-439-21
; Sequence 21, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439-21

Query Match      31.68; Score 1775.5; DB 12; Length 868;
Best Local Similarity 50.38; Pred. No. 7.7e-144;
Matches 335; Conservative 122; Mismatches 186; Indels 23; Gaps 7;

QY      31 REVSSVQLSEESTFYLCPPPVGVSTVIRLEPPKCPKPRKATEWGEGIAILFKENISPY 90
Db      41 REAIKHSQDAETPTFFVCPPPGTSTVIRLEPTCPDYHLGNFTGEGIAVTKENIAAY 100
QY      91 KFKVTLYKNIQTTWTGTYTQIRNRYDRTPTVSLIEETDLDIGKGRCSKARYLRNN 150
```

```
Db      101 KFKATVYKDVIVSTAWAGSSYTQITNRYADRVPIPVSEITDTIDKFGKSSKATYVRNN 160
QY      151 VVVEAFORDAGEKOVLLKPSKFNPESRAWHTTNETYTVMGSPWLYRTGTSTVNCIVEEMD 210
Db      161 HKVEAFNEKDPQDPLTIASKYNSVGSRAWHTTNDTYMVGTPGTGYRTGTSTVNCILIEBE 220
QY      211 ARSVEPKSYFAMANGDIANISPFYGLSPPEAAEPMPQDNFKQDLSYFSMDLDRKKA 270
Db      221 ARSIFPYDSGLSGTDIIYMSPFEGLR-DGAYREHSNTYAMDRFHQFEGYRQDLDTRAIL 279
QY      271 SLPKVRNPLIISHTVGHDWAPKPTRVCSMTKWKEVTEMLRATVNGRYFRFMARELSAF 330
Db      280 E-PAARNFLVPLHTVGMNWKPKRTEVCSLVKWEVEDVVRDEYAHNFRFTMKLTSTTFI 338
QY      331 SNTTFEDFNRIILGQCIKREAEAEQIFRTKYNDSHVKVGHVQYFELALGQFIVAYOPVL 390
Db      339 SETNEFNLIQHLSCVKEEARAIINRIYTRYNSSHVTRTDIQTILARGGFVVVFQPLL 398
QY      391 SKSLAHMYLRELMRDRNTDEMLDLVNKNHAIYKKKNATSLSLRLRRDIRNAPNRKITLDDTT 450
Db      399 SNSLARLYLQELVRE-----NTNHS-PQKHPTNTRTSRRSV-----PVELRANR 441
QY      451 AIKSTSSVQFAMLOFLYDHIQTHINDMFSRIATACELONRELVLWHEGIIKINFSATASA 510
Db      442 TITTISSVEFAMLOFTYDHIQHEVNMELARISSSMCQONRERALSGLFPIINFSALAST 501
QY      511 TIGRRVAAKMLGDVAAYSSCTAIDAES-VTLQNSMRVITSTNTCVSRPLVLSYGENOGN 569
Db      502 ILDQVRKARIILGDVISVSNCPGLGSDTIFILQNSMRVSGSTRCYSRPLISIVSLNGSGT 561
QY      570 IQGQIGENNELLPTELEAVEPCSANHRYFLFGSYALFENFYKMWDAADIQIASTFVE 629
Db      562 VEGQLGTDNELIMSRLLEPCVANKRYFLFGHHYVYEDYRYVREITAVHDVGMISTYVD 621
QY      630 LNLTLLEDREITLPLSVYTKELRDVGLDVAEVARRNOLHELKPYDINKVIEVDNTYA-- 687
Db      622 LNLTLKDRFMPLOQVYTRDELDTGLDYSEIQRRNQMHSLRFYDIDKVVQYDSGTAIM 681
QY      688 -GLOEF 692
Db      682 QGMAQF 687
```

```
RESULT 10
US-09-301-390-21
; Sequence 21, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-301-390-21

Query Match 31.6%; Score 1775.5; DB 17; Length 868;
Best Local Similarity 50.3%; Pred. No. 7.7e-144;
Matches 335; Conservative 122; Mismatches 186; Indels 23; Gaps 7;

Qy 31 REVYSSVOLSEESTFVLCPPPGVSTVIRLEPPKCPKPRKATEWGEGIALFKENISPY 90
Db 41 REAIHKSDAETKFTFYVCPPTGTVLRLEPTRTCPDYHLGKNFTGIAVYKENIAAY 100

Qy 91 KFKVTLXYKNIQIOTTTGTTRQITNRYTDRTPVSEIEITDLDGKRCSSKARYLRNN 150
Db 101 KFKATVYKDVIVSTAWAGSSYQITNRYADRPVPIVSEITDIDKFKCCKSKATYVRNN 160

Qy 151 VYVAFORDAGEKOVLLKPSFNPESRAWHTTNETYVMGSPWYRTGTSVNCIVEMD 210
Db 161 HKVEAFNEBKRPQDMPLIASKYNSVGSKAWHTTNDYVAVAGTPGTGTSVNCILIEVE 220

Qy 211 ARSVFPYSYFAMANGDANISPFYGLSPPEAAAPMGYPQDNFKOLDSYFSDLDKRRKA 270
Db 221 ARSIFPYDSFGISGDIYIYSPFFGLR-DGAYREHSNYAMDRFHQFEGYRQDLDTALL 279

Qy 271 SLPVKRNFLITSHTFVGWDMAPKTRVCSMTKWEVTEMLRATVNGRYRFAWELSATFI 330
Db 280 E-PAARNFLVTPHLTVGNWPKRTEVCSLVKWEVEDVDVDEYAHNFRFTMKLTSTFI 338

Qy 331 SNTTEFPDNRILQGCCKREAAAEIQIFRTKYNDSHVKVGHVQYFLALGGFIVAYQPV 390
Db 339 SETNEFNLNQIHLSCQVKEEARAIINRIYTRYNSSHVHTGDIQTYLARGGFVVFQPLL 398

Qy 391 SKSLAHMYLRELMDRNTDMDLVNKNHAIYKKNATSLRLRDRINAPNRKTYLDDTT 450
Db 399 SNSLARLYQLVLVE-----NTNHS-PQKHPTNTRSRRSV-----PVELRANR 441

Qy 451 AIKSTSSQFAMQLFDYDHIQTHINDMFRIATANCELRNRELVLWHEGKINPSATASA 510
Db 442 TITTTSSVEFAMQLFTYDHIQEHVNEMLARISSCQQLONRERALLWSGLFPINPSALAST 501

Qy 511 TGLRRVAAKMGIDVAASSCATAIDAES-VTLQNSMRVITSTNTCYSRPLVLFSGYENQGN 569
Db 502 ILDQVRKARILGDVIVSNCELSGDRILLQNSMRVSGSTTRYCYSRPLISVLSNGST 561

Qy 570 IQGQGENNELPTLEAVEPCSAHRRYFLFGSGYALFENFYVKWDAADIQIASTFVE 629
Db 562 VEGQLGTDNELMSRDLLEPCVANHKKRYFLGHHVYVYEDYRYVREIAVHDVGMISTVVD 621

Qy 630 LNLILLEDRELPLSVYTKELRDVGLDYAEVARRNQLHKLKYDINKVLEVTNTA-- 687
Db 622 LNLILKDRFMPQLVYTRDELRTGLLDYSEIQRRNQMHSLRFYDIDKVVQYDSGTAIM 681

Qy 688 -GLOEF 692
Db 682 QGMAQF 687

RESULT 11
US-09-338-326-21
; Sequence 21, Application US/09338326
; GENERAL INFORMATION:

```

```

; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-338-326-21

Query Match 31.6%; Score 1775.5; DB 17; Length 868;
Best Local Similarity 50.3%; Pred. No. 7.7e-144;
Matches 335; Conservative 122; Mismatches 186; Indels 23; Gaps 7;

Qy 31 REVYSSVOLSEESTFVLCPPPGVSTVIRLEPPKCPKPRKATEWGEGIALFKENISPY 90
Db 41 REAIHKSDAETKFTFYVCPPTGTVLRLEPTRTCPDYHLGKNFTGIAVYKENIAAY 100

Qy 91 KFKVTLXYKNIQIOTTTGTTRQITNRYTDRTPVSEIEITDLDGKRCSSKARYLRNN 150
Db 101 KFKATVYKDVIVSTAWAGSSYQITNRYADRPVPIVSEITDIDKFKCCKSKATYVRNN 160

Qy 151 VYVAFORDAGEKOVLLKPSFNPESRAWHTTNETYVMGSPWYRTGTSVNCIVEMD 210
Db 161 HKVEAFNEBKRPQDMPLIASKYNSVGSKAWHTTNDYVAVAGTPGTGTSVNCILIEVE 220

Qy 211 ARSVFPYSYFAMANGDANISPFYGLSPPEAAAPMGYPQDNFKOLDSYFSDLDKRRKA 270
Db 221 ARSIFPYDSFGISGDIYIYSPFFGLR-DGAYREHSNYAMDRFHQFEGYRQDLDTALL 279

Qy 271 SLPVKRNFLITSHTFVGWDMAPKTRVCSMTKWEVTEMLRATVNGRYRFAWELSATFI 330
Db 280 E-PAARNFLVTPHLTVGNWPKRTEVCSLVKWEVEDVDVDEYAHNFRFTMKLTSTFI 338

Qy 331 SNTTEFPDNRILQGCCKREAAAEIQIFRTKYNDSHVKVGHVQYFLALGGFIVAYQPV 390
Db 339 SETNEFNLNQIHLSCQVKEEARAIINRIYTRYNSSHVHTGDIQTYLARGGFVVFQPLL 398

Qy 391 SKSLAHMYLRELMDRNTDMDLVNKNHAIYKKNATSLRLRDRINAPNRKTYLDDTT 450
Db 399 SNSLARLYQLVLVE-----NTNHS-PQKHPTNTRSRRSV-----PVELRANR 441

```


552 VNEMLSRIATAWCTLQNEHRLWTETTLKLNPGGVYSMALERRYSARLLDGA VAVTCVNI 611
 QY 534 DAESTVTLQNSMRVITSTWCTYSRPLVLFSGYENOGNIQOLGENNELLP TLEAVEPCSAN 593
 Db 612 SSGHYVQNSMRVTCSSITCSRPLVSFRALNDEYEGQLGENNELLVKRLIEPCTVN 671
 QY 594 HRRYLFSGGYALFENYFVKWDAADIQIASTAEVNLTLLEDRETLPLSVYTKBELRD 653
 Db 672 NKRYEFGADYVYFEDYAVRKVPLSETELISAVV-IKSTLLEDREPLH-SSYTRAELED 729
 QY 654 VGVLDYAEVARNQLHELKPFVDINKVIEVDTN-----YAGLQEFCC----- 694
 Db 730 TGFPEYSEIQRRNQLHALKFYDIDSVIRVDNNLVIMRGMANFQGLGDVGAGFGKVV LGA 789
 QY 695 -----MSITKKDANPNNGOTOLEAARMELTDLINA-----KAM---TLAS 731
 Db 790 ASAVISTVSGVSSFLNNPFGALVGLLILAGIAAFLAYRYSRLRANPKMALYPVTRN 849
 QY 732 LDYAKIEASLSSAYSEAEVTNNNLNATLEQLKMAK-TNLESAINQANTDKTTTFDNEHPN 790
 Db 850 LKQATAKSPASTAGGSDPGVDDEEKLQAREMIKYSLSVAMEQ-QEHKAMKKNGPA 908
 QY 791 LIVEAY---KALKT---TLEQRATNLEGLSSATYNOI 820
 Db 909 ILTSLTNMAULRRRGPKYQRLNNLNDSGDDTENV 943

 RESULT 14
 US-09-521-738-14
 ; Sequence 14, Application US/09521738
 ; GENERAL INFORMATION:
 ; APPLICANT: Jensen, Wayne A.
 ; APPLICANT: Lappin, Michael R.
 ; APPLICANT: Rosen, David K.
 ; APPLICANT: Andrews, Janet S.
 ; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
 ; FILE REFERENCE: DI-9
 ; CURRENT APPLICATION NUMBER: US/09/521,738
 ; CURRENT FILING DATE: 2000-03-09
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 943
 ; TYPE: PR1
 ; ORGANISM: Feline herpesvirus 1
 US-09-521-738-14

 Query Match 31.0%; Score 1743; DB 19; Length 943;
 Best Local Similarity 40.3%; Pred. No. 5.9e-141;
 Matches 377; Conservative 161; Mismatches 267; Indels 130; Gaps 21

 QY 3 YFRNCIF-----FLVILY-----GTNSSP 23
 Db 22 YFRCQFFP SLLGIAATGRHGSSGLTRLARYVSFIWVLFVGRPVVEGSGSTSEQ 81
 QY 24 STQNTSREV-----VSSVOLSEEE---STFVLCPPVPVGSIVRLRPP 63
 Db 82 PRRTVATPEVGVHQNQIQIPICRYEALRASQIEANGSTFWCMPPSGSTVWRLEPP 141
 QY 64 RKCPEPRKATEWEGEIALFKENISPKFKVLYYKNIQTWTGTYTQINRYTDR 123
 Db 142 RACPDYKLGKNFTGEGIAVIFENIAPYKFKANIYKNIIMTWGSSSYAVTTNRYDRV 201
 QY 124 PVSIEITDLIDKGRCSKARYLRNNVYVFAFDAGEQVLLKPSKFNWTPESAWHTT 183
 Db 202 PVKQVEITDLIDRRGMLSKADYVRNNYQFTAFDRDEDPRELPLKPPSTLSRVGWH-T 260
 QY 184 NETYTVWGSWPYIRGTSYNCIVEMDARSVPFYSYFAMANGDIANISPPYGLSPPEAA 243
 Db 261 NETYTKVILLDFPHHSGTSVNCIVEDARSYPYDSPAISTGVHISPPFFGLR-DGAHV 319
 QY 244 EPMGYPODNFKOLDYSYFSDMDLKRKASLPVKRNFLLTSHFTVGVGDWADKTTTVC-SWTKW 303

```
Db 320 EHTSSRRFQIEGYIDLDTDYGA-PSVRNELEPHVTVAWNTPKSGRVCITLAKW 378
QY 304 KEVTEMLRATVNGRYRFMARELSATISNTTEFDNRIILGQCIRKREAAIEQIFRTKY 363
Db 379 REIDMLPMNI-GSYRFTAKTISAFISNTSQFEINRIRLGCAKKAEEAIDRIYKSY 437
QY 364 NDHSVHVGVQVFLALGFIYAVQVLSKSLAHMYLRELMRDNRDEMDLDLNNKHAYK 423
Db 438 SKHTQTGTLETYLRAGGFIAFRPMISNELAKLYINELARSNT--VVDL-----SALLN 491
QY 424 KNATSLSLRRDIRNAPRKI-----TLDDTTAKTSSVQFAMQLQFLYDHIQTH 473
Db 492 PSGETVQTRSPVPSNQHRSRSTIEGGIETVNNASLLKTTSSVEFAMLOFAYDIQAH 551
QY 474 INDMFSRIATAWCELQNLRELWLBHGIKINPSATISATILGRVAAKMLGDVAASCTAI 533
Db 552 VNEMLSRIATAWCTIQNLREHLWTETLKLNGGVVSMALERRVSARLLGDVAVTCQVNI 611
QY 534 DAEVTLQNSMRVITSTWTCYSRPLVLSYGENQNIQOLGENNELPTLEAVEPCSAN 593
Db 612 SSGHYIIONSRVTSSTCYSRPLVSPRALNDSEYIEGQNLGENNELVERKLIEPCTVN 671
QY 594 HRRYFLGSGVALPENYFVMDVDAADIOIASTFVELNLTLLEDEILPLSYTKEELRD 653
Db 672 NKRYFKGADYVYFEDYAVRKVPLSEIELISAYV-IKSTLLEDFELH-SSYTRAELED 729
QY 654 VGVLDYAEVARNQLHELKFDYDINKVIEVDN-----YAGLOBFGC-----694
Db 730 TGPFDYSEIQRNQLHALKFYDIDSIVRVDNNLVIRMGMANFFQGLGDVGAGFGKVLGA 789
QY 695 -----MSITKTDANPNNGQTOLEARMELTDLINA-----KAM---TLAS 731
Db 790 ASAVISTVSGVSSFNNPFGALVGLLLIAGIVAAFLAYRYSIRLRNPMKALYVPTTRN 849
QY 732 LDYAKIEASLSAYSEAYETVNNLNATLEQKMAK-TNLESAINQANTDKTTFDNEHPN 790
Db 850 LKQTAESPASTAGSDGPDVDFDEKLMQAREMIKYMSLYSAMQ-QEHKAMKKNKGA 908
QY 791 LVEAY--KALKT--TLQBATNLEGLSSTAYNQI 820
Db 909 ILTSLTNMLRRRPGKYQRLNLDSDGDDTETNLV 943
```

RESULT 15

```
US-08-541-878-6
; Sequence 6, Application US/08541878
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Black, Darla
; APPLICANT: Scinicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
; TITLE OF INVENTION: Virus Genes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cox & Smith Incorporated
; STREET: 112 East Pecan Street, Suite 2000
; CITY: San Antonio
; STATE: Texas
; COUNTRY: USA
; ZIP: 78205
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,878
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/042,747
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haymond, W. Bradley
; REGISTRATION NUMBER: 35186
; REFERENCE/DOCKET NUMBER: S-0072.179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 210-554-5500
; TELEFAX: 210-226-8395
; TELEX: 767609
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 891 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-541-878-6

Query Match 31.0%; Score 1739.5; DB 9; Length 891;
Best Local Similarity 49.0%; Pred. No. 1.1e-140;
Matches 332; Conservative 103; Mismatches 216; Indels 27; Gaps 3;

QY 18 GTNSSPSTQNTSREVSVSVLSEESTFYLCPPPVGVSTVIRLEPPKCPKPRKATEGE 77
Db 72 GTNASVEAGHATLRNLRDIALDGDATFYVCPPTGATVVGFEQPCRPAPHCQNTKE 131
QY 78 GIALFKENISPYKFKVTLXYKNIITQTTTGTGYRQITNRYTDRTPYSIEBITDLIDGK 137
Db 132 GIATVIFKENIAPYKFKATMYKDVTVSQVWFCHRYSQFMGIFEDRAPVPFEEVDIKINAR 191
QY 138 GRCSKARYLNNTVVEAFDRDAGEKQVLLAKPSKENTPESRAWHHTNETYTVWGSFWIYR 197
Db 192 GVCSTAKYVRNNMESTAFHDDDDSDMKLPKAKAATRTSRGWHHTDLUKYNSRTEAFHR 251
QY 198 TGTSYNCIVEEMDARSFPYSYFAMANGDIANISPFYGLSPPEAAEPMPQDQNFOLD 257
Db 252 YGTTVNCIVEEVEARSVTPYDEFVLATGDFVYMSFFYGR-DGAHAETAATAADRFRQVD 310
QY 258 SYFSMDLDKRKASLPVKRNFILITSHFVGVNDWAPKTRVCSMTKWKEVTEMLRATVNGR 317
Db 311 GYERDLSTGRRASTPATRNLLTTPKFTVGVNDWAPKRPSCVCTLTWKQVDEMLRAEYGPS 370
QY 318 YRFMARELSATFISNTTEFDNRIILGQCIRKREAAIEQIFRTKYNDSHVKVGHVQVFL 377
Db 371 FRFSSALSTFTTNRTEYALSRVDLGCVGREAREADVDRIFLRRYNGTHVKVGQVQYIL 430
QY 378 ALGGFIVAYQVPLSKSLAHMYLRELMRDNRDEMDLDLVNNKHAIYKKNATSLSLRRDIR 437
Db 431 ATAGFLIAYQPLLSNGLVELYVRELLREQGRP-----GDAA 467
QY 438 NAPHKTIILDDTTAKTSSVQFAMQLQFLYDHIQTHINDMFSRIATANCELQNLRELVLWH 497
Db 468 ATPKPSADPPDVERIKTTSSEFARLQFTYDHIQHVNDMLGRIAIANCELQNLHELTLWN 527
QY 498 EGKINPSATASATILGRVAAKMLGDVAASCTAIDAESVTQLONSMEVISTNTCYSRP 557
Db 528 EARLKNPNAIASATVGRVRSARMLGDVMAVSTCVPTVPDINVMQNSMRVAPRPGTCYRSP 587
QY 558 LVLFYSYGENOQNIQOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFVMDV 617
Db 588 LVSFYREBGGPLVEGQLGEDNEIRLEDALEPCTVGHRRYFTFGAGYVYFEDYAYSHQLG 647
QY 618 AADIQIASTFVELNLTLLEDEILPLSYTKEELDVGVLDAEYAEVARNQLHELKFDYDIN 677
Db 648 RADVTTVSTFTNLNLTMLEDEHFEVPLVEYTRQETKDSGLDLYTEVQRNQLHALRFADID 707
QY 678 KVEVDN---YAGLOEF 692
Db 708 TVIKADHAPLFAGLYSF 725
```

Search completed: October 8, 2003, 17:24:58
Job time : 438.972 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:07:16 : Search time 62.4796 Seconds
(without alignments)
2800.682 Million cell updates/sec

Title: US-09-147-052-4
Perfect score: 5619
Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNENADKIPGRRPGTFL 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5619	100.0	456	9	US-09-147-052-4
2	2015	35.9	456	9	US-09-147-052-2
3	2008	35.7	456	12	US-09-901-572A-3
4	1929.5	34.3	384	15	US-10-131-591A-12
5	1777	31.6	357	15	US-10-131-591A-23
6	1736	30.9	904	9	US-09-894-998-18
7	1736	30.9	904	12	US-10-200-562-18
8	1736	30.9	904	12	US-10-237-551-18
9	1736	30.9	904	15	US-10-121-988-18
10	778.5	13.9	907	15	US-10-223-538-2
11	758.5	13.5	830	14	US-10-055-364-45
12	754	13.4	865	14	US-10-055-364-24
13	750	13.3	793	14	US-10-055-364-55
14	738.5	13.1	844	14	US-10-055-364-40
15	725.5	12.9	808	14	US-10-055-364-38

16	719	12.8	359	14	US-10-055-364-48	Sequence 48, Appl
17	716	12.7	823	14	US-10-055-364-37	Sequence 37, Appl
18	709.5	12.6	829	14	US-10-055-364-46	Sequence 46, Appl
19	704	12.5	831	14	US-10-055-364-39	Sequence 39, Appl
20	700.5	12.5	818	14	US-10-055-364-44	Sequence 44, Appl
21	695	12.4	791	14	US-10-055-364-41	Sequence 41, Appl
22	695	12.4	792	14	US-10-055-364-42	Sequence 42, Appl
23	679.5	12.1	824	14	US-10-055-364-43	Sequence 43, Appl
24	595.5	10.6	206	14	US-10-127-733-2	Sequence 2, Appl1
25	343.5	6.1	195	14	US-10-055-364-2	Sequence 2, Appl1
26	314	5.6	62	15	US-10-131-591A-5	Sequence 6, Appl1
27	302	5.4	62	15	US-10-131-591A-6	Sequence 5, Appl1
28	216.5	3.9	2659	12	US-10-311-879-28	Sequence 28, Appl
29	188.5	3.4	2478	9	US-09-815-242-5816	Sequence 5816, Ap
30	188.5	3.4	2478	9	US-09-815-242-12967	Sequence 12967, A
31	187.5	3.3	2368	9	US-09-815-242-5635	Sequence 5635, Ap
32	187.5	3.3	2368	9	US-09-815-242-12389	Sequence 12389, A
33	186.5	3.3	2086	9	US-09-815-242-5639	Sequence 5639, Ap
34	186.5	3.3	5795	9	US-09-815-242-12610	Sequence 12610, A
35	186.5	3.3	6281	9	US-09-815-242-12996	Sequence 12996, A
36	180	3.2	1029	9	US-09-815-242-5885	Sequence 5885, Ap
37	180	3.2	1048	9	US-09-815-242-13083	Sequence 13083, A
38	177.5	3.2	2434	9	US-09-815-242-5835	Sequence 5835, Ap
39	176	3.1	3712	14	US-10-108-605-103	Sequence 103, App
40	174	3.1	2437	9	US-09-815-242-5834	Sequence 5834, Ap
41	168.5	3.0	1433	10	US-09-801-368-60	Sequence 60, Appl
42	164	2.9	1069	15	US-10-146-473-77	Sequence 77, Appl
43	163.5	2.9	1220	12	US-10-193-764-28	Sequence 28, Appl
44	163.5	2.9	1226	12	US-10-193-764-26	Sequence 26, Appl
45	163.5	2.9	1948	12	US-10-032-585-7611	Sequence 7611, Ap

ALIGNMENTS

RESULT 1

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

Query Match	100.0%;	Score 5619;	DB 9;	Length 1086;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1086;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	MHYFRNCIFFLIVLYGTNSPSTONVTSREVSVSSVOLSEESTFYLCPPPGVSVIRL	60	
Db	1	MHYFRNCIFFLIVLYGTNSPSTONVTSREVSVSSVOLSEESTFYLCPPPGVSVIRL	60	
QY	61	EPRKCPERKATWEGEGTALFKENISPKFKVLYYKNIIOTTWTGTYQINRYT	120	
Db	61	EPRKCPERKATWEGEGTALFKENISPKFKVLYYKNIIOTTWTGTYQINRYT	120	
QY	121	DRTPVSEETDLDIGKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPKFKNTPE\$RAW	180	

```

121 DRTPVSEIETDLDGKRCSSKARYLRNNVVEAFDRDAGEQVLLKPSKENTPESRAW 180
181 HTTNETYVWSPWIRYGTSTVNCIIVEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
181 HTTNETYVWSPWIRYGTSTVNCIIVEMDARSVPYSYFAMANGDIANISPFYGLSPPE 240
241 AAAEPMGYPQDNFKOLDSYFSDMDLKRKASIPVKRNELITSHFTVGVGDWAPKTRVCSM 300
241 AAAEPMGYPQDNFKOLDSYFSDMDLKRKASIPVKRNELITSHFTVGVGDWAPKTRVCSM 300
301 TKWKEVTEMLRATNGRYRFMARLSATFISNTTEFDPNRIILGQCIRKREAAAEQIFR 360
301 TKWKEVTEMLRATNGRYRFMARLSATFISNTTEFDPNRIILGQCIRKREAAAEQIFR 360
361 TKYNDSHVKGHVQYFALGQFIVAYQPVLSKSLAHMYLRELNRDNRDEMLDLVNNKHA 420
361 TKYNDSHVKGHVQYFALGQFIVAYQPVLSKSLAHMYLRELNRDNRDEMLDLVNNKHA 420
421 IYKKNATLSRLRDIRNAPNRKITLDDTTAKTSSVQFAMLOFLYDHIQTHINDMFSR 480
421 IYKKNATLSRLRDIRNAPNRKITLDDTTAKTSSVQFAMLOFLYDHIQTHINDMFSR 480
481 IATAWCELQNLRELVLWHEGIRKINPSATASATLGRVAAKMDGVAASVSSCTAIDAESVTL 540
481 IATAWCELQNLRELVLWHEGIRKINPSATASATLGRVAAKMDGVAASVSSCTAIDAESVTL 540
541 QNSMRVITSTNCTSRPLVLFSGENQGNIOGQGENNELLPTEAVEPCSANHRRYPLF 600
541 QNSMRVITSTNCTSRPLVLFSGENQGNIOGQGENNELLPTEAVEPCSANHRRYPLF 600
601 GSGYALFENYFVKMVAADIIQIASTFVELNLTLLEDEILPLSYTKEELRDVGVLDA 660
601 GSGYALFENYFVKMVAADIIQIASTFVELNLTLLEDEILPLSYTKEELRDVGVLDA 660
661 EVARRNQLHELFYDINKVIEVDNYAGLQEFQCMSTKKDANPNNGQTOLEAARMELTD 720
661 EVARRNQLHELFYDINKVIEVDNYAGLQEFQCMSTKKDANPNNGQTOLEAARMELTD 720
721 LINAKAMTLASLDYAKTEASLSAYSEAEVTVNNNLNATLQKMAKTNLESAINQANTD 780
721 LINAKAMTLASLDYAKTEASLSAYSEAEVTVNNNLNATLQKMAKTNLESAINQANTD 780
781 KTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQIRNNLVLDLYNKASSLITKTL 840
781 KTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYNQIRNNLVLDLYNKASSLITKTL 840
841 PLNGGTLDSNEITTANKNINNTLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNA 900
841 PLNGGTLDSNEITTANKNINNTLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNA 900
901 NVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSWIYSLAGNTK 960
901 NVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSWIYSLAGNTK 960
961 YQSFNSYGPSTGYLYPPYKLVKAADANNVGLQYKLNNGNVQVVEFATSTSANNTTANPT 1020
961 YQSFNSYGPSTGYLYPPYKLVKAADANNVGLQYKLNNGNVQVVEFATSTSANNTTANPT 1020
1021 PAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADKIPGYR 1080
1021 PAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNENNADKIPGYR 1080
1081 RPKTEFL 1086
1081 RPKTEFL 1086

```

RESULT 2

```

US-09-147-052-2
; Sequence 2, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:

```

```

; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

```

```

Query Match 35.9%; Score 2015; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 5.2e-150;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 691 EFGCMSITTKDANPNNGQTOLEAARMELTDLINAKAMTLASLDYAKTEASLSAYSEAE 750
Db 61 EFGCMSITTKDANPNNGQTOLEAARMELTDLINAKAMTLASLDYAKTEASLSAYSEAE 120
QY 751 TVNNNLNATLQKMAKTNLESAINQANTDITFDNEHPNLVEAYKALKTTLEQATNLE 810
Db 121 TVNNNLNATLQKMAKTNLESAINQANTDITFDNEHPNLVEAYKALKTTLEQATNLE 180
QY 811 GLSSTAYNQIRNNLVLDLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQ 870
Db 181 GLSSTAYNQIRNNLVLDLYNKASSLITKTLPLNGGTLDSNEITTANKNINNTLSTINEQ 240
QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVN 930
Db 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVN 300
QY 931 GDEPSSRLANTNSITDVSWIYSLAGNTKYOFSFSGYLYPPYKLVKAADANNV 990
Db 301 GDEPSSRLANTNSITDVSWIYSLAGNTKYOFSFSGYLYPPYKLVKAADANNV 360
QY 991 GLQYKLNNGNVQVVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
Db 361 GLQYKLNNGNVQVVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
QY 1051 EGNMKNKVPAMIGNIYLSNENNADKIPGYRPGTFL 1086
Db 421 EGNMKNKVPAMIGNIYLSNENNADKIPGYRPGTFL 456

```

RESULT 3

```

US-09-901-572A-3
; Sequence 3, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

```


Query Match 35.7%; Score 2008; DB 12; Length 456;
Best Local Similarity 99.5%; Pred. No. 1.9e-149;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 691 EFGCSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE 750
DB 61 EFGCSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE 120

QY 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 810
DB 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180

QY 811 GLSSTAYNQIRNNLDVLYNKASLITKTTLDPLNGGTLTLDSEITTTANKNINNTLSINEQ 870
DB 181 GLSSTAYNQIRNNLDVLYNKASLITKTTLDPLNGGTLTLDSEITTTANKNINNTLSINEQ 240

QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARTVWN 930
DB 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARTVWN 300

QY 931 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 990
DB 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 360

QY 991 GLOYKLNNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
DB 361 GLOYKLNNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420

QY 1051 EGNMKNVAPMIGNIYLSNENNAKIPGYRRPGTFL 1086
DB 421 EGNMKNVAPMIGNIYLSNENNAKIPGYRRPGTFL 456

RESULT 4

US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Query Match 34.3%; Score 1929.5; DB 15; Length 384;
Best Local Similarity 99.2%; Pred. No. 2.1e-143;
Matches 382; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 691 EFGCSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE 750
DB 1 EF-CMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAE 59

QY 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 810
DB 60 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 119

QY 811 GLSSTAYNQIRNNLDVLYNKASLITKTTLDPLNGGTLTLDSEITTTANKNINNTLSINEQ 870
DB 120 GLSSTAYNQIRNNLDVLYNKASLITKTTLDPLNGGTLTLDSEITTTANKNINNTLSINEQ 179

QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARTVWN 930
DB 180 KTNADALSNSFIKKVIONNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARTVWN 239

QY 931 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 990
DB 240 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 299

QY 991 GLOYKLNNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
DB 300 GLOYKLNNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 359

QY 1051 EGNMKNVAPMIGNIYLSNENNAK 1075
DB 360 EGNMKNVAPMIGNIYLSNENNAK 384

RESULT 5

US-10-131-591A-23
; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified Ttl-1 portion (downstream of BglI) of
US-10-131-591A-23

Query Match 31.6%; Score 1777; DB 15; Length 357;
Best Local Similarity 98.6%; Pred. No. 1.9e-131;
Matches 352; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 719 TDLINAKAMTSLASQDYAKIEASLSAYSEAEETVNNLNATLEQLKMAKTNLESAINQAN 778
DB 1 TDLINAKAMTSLASQDYAKIEASLSAYSEAEETVNNLNATLEQLKMAKTNLESAINQAN 60

QY 779 TDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQIRNNLDVLYNKASLITKT 838
DB 61 TDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQIRNNLDVLYNKASLITKT 120

QY 839 LDPNGGTLTLDSEITTTANKNINNTLSINEQNTADALSNSFIKKVIONNEQSFVGTFT 898
DB 121 LDPNGGTLTLDSEITTTANKNINNTLSINEQNTADALSNSFIKKVIONNEQSFVGTFT 180

QY 899 NANVQPSNYSFVAFSADVTPVNYKYARTVWNGDEPSSRILANTNSITDVSIIYSLAGTN 958
DB 181 NANVQPSNYSFVAFSADVTPVNYKYARTVWNGDEPSSRILANTNSITDVSIIYSLAGTN 240

QY 959 TKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNVQOQVEFATSTSANNTTAN 1018
DB 241 TKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNVQOQVEFATSTSAQOTTAN 300

QY 1019 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNAK 1075
DB 301 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNAK 357

RESULT 6

US-09-894-998-18
; Sequence 18, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894.998
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-09-894-998-18

Query Match 30.9%; Score 1736; DB 9; Length 904;
Best Local Similarity 43.0%; Pred. No. 1.5e-127;
Matches 363; Conservative 142; Mismatches 269; Indels 70; Gaps 16;

Qy 20 NSSPSTQNTVSRVSVSSVQLSEESTFVLCPPPVGVSTVIRLEPPKCPPEPRKATEWEGE 79
Db 82 NATVAAGHATLRAHLREIKYENADAQFYVCPPTGATVQVQEPQPRCPTRPEGQNYTEGI 141

Qy 80 AILFKNISPKFKVLYYKNIOTTWTGTYRQITNRYDTRTPVSEIEITLIDGKGR 139
Db 142 AVFKEKNIAPYKFKATYKDVTVSQVMFGHRYSQFMGIFEDRAPVPFEEVIDKINTKGV 201

Qy 140 CSSKARYLRNNVYVEAFDRAGEKQVLLKPSKFTWPESRAWHHTTNEYTVWGSPIWYRTG 199
Db 202 CRSTAKYVRNNMETAFHRDDHETDMELPKAVATRTSRGWHHTTDLKYNPSRVEAFHRYG 261

Qy 200 TSVNCIVEEMDARSVPFYSYFAMANGDIANTISPPYGLSPPPAAAEPMGYPODNFKQLDSY 259
Db 262 TTVNCIVEEVDARSVPYDEPLATGFVYMSPPGYR-EGSHTHTSYAADRFKQVDGF 320

Qy 260 FSDMLDRRKASLPVKRNFLLTSHFTVGVWDWAPKTRTVCSMTKWKVEYEMLRATVNGRYR 319
Db 321 YARDLTTKARATSPTRNLLTTPKFTVAVDWVPRPAVCTMTKQWEDMLRAEYGGSGFR 380

Qy 320 FMARELSATFISNTTEPDNRIILGOCIKREAAEAAIEQIFTKYNDSHVKVGHVQYFLAL 379
Db 381 FSSDAISTFTTNLTTEYSLSRVDLGDGICGRDAREIDRMFAKNATHYKVGQPYLYAT 440

Qy 380 GGFTVAYQPVLSKSLAHMYLRELNRDNTDMLDNVNNKHAIYKKNATLSRLRDRINA 439
Db 441 GGFLIAYQPLLSNTLAELYVREYMQ-----DRK-----PRNATP-----APLREA 482

Qy 440 PNRIITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSRIATAWELQNLRELVLNHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHQIRHVNDMLGRIVAWCELQNLHETLWNEA 539

Qy 500 IKINPSATASATLGRVVAAMGLDVAVSSCTAIDAESVTLONSMRVITTSNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMGLDGMVAVSTCVPAQDNVIVQNSMRVSSRPGTCYSRPLV 599

Qy 560 LFSYGENQGN-IOGOLGENNELPTLEAVEPCSAHRRYFLFGSGYALFENYFVKMVA 618
Db 600 SFRY-EDQGPLIEGOLGENNELRLTRDALEPCTVGHRRYFFGGYVYFEEYAYSHQLSR 658

Qy 619 ADIOIASTFVNLTLLEDEILPLSVYTKTELSDVGLDYAEVARNQLHELKFDYDINK 678
Db 659 ADVTTVSTFIDLNTMLDEHFEVPLEVYTRHEIKDGLDYTEVQRNQLHDLRFADIDT 718

Qy 679 VIEVDTN---YAGIQEF-----GCMSTIKKDNPNNGQTOLEARMELTDLI 722
Db 719 VIRADANAMEFGLCAFEFGMDLGRAVGKVMGVGVGVYSAVSGVSSF-----MSNPF 772

Qy 723 NAKAMTLASLDQYAKIEASISSAYSEATV-----NNLNLA-----TLBQLMAKNTLSEAI 774
Db 773 GALAVGLLVL-----AGLVAFAFRYVLQLRPNPKALYPLTTKELKTSDDPGVGGE 825

Qy 775 NQANTDKTTFDNEHPLNVEA-----YKALKTLE--QBATNLEGLSTAYNQIRNLVDL 827
Db 826 GEEGAEGGGFD--BAKLAEAREMIRYMALVSAMERTEHKARKKGTSAALLSSKVTNMLVRK 883
```

```
Qy 828 YNKA 831
Db 884 RNKA 887

RESULT 7
US-10-200-562-18
; Sequence 18, Application US/10200562
; Publication No. US20030165819A1
; GENERAL INFORMATION:
; APPLICANT: McGowen, Patrick
; APPLICANT: Hosken, Nancy A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C2
; CURRENT APPLICATION NUMBER: US/10/200.562
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-200-562-18

Query Match 30.9%; Score 1736; DB 12; Length 904;
Best Local Similarity 43.0%; Pred. No. 1.5e-127;
Matches 363; Conservative 142; Mismatches 269; Indels 70; Gaps 16;

Qy 20 NSSPSTQNTVSRVSVSSVQLSEESTFVLCPPPVGVSTVIRLEPPKCPPEPRKATEWEGE 79
Db 82 NATVAAGHATLRAHLREIKYENADAQFYVCPPTGATVQVQEPQPRCPTRPEGQNYTEGI 141

Qy 80 AILFKNISPKFKVLYYKNIOTTWTGTYRQITNRYDTRTPVSEIEITLIDGKGR 139
Db 142 AVFKEKNIAPYKFKATYKDVTVSQVMFGHRYSQFMGIFEDRAPVPFEEVIDKINTKGV 201

Qy 140 CSSKARYLRNNVYVEAFDRAGEKQVLLKPSKFTWPESRAWHHTTNEYTVWGSPIWYRTG 199
Db 202 CRSTAKYVRNNMETAFHRDDHETDMELPKAVATRTSRGWHHTTDLKYNPSRVEAFHRYG 261

Qy 200 TSVNCIVEEMDARSVPFYSYFAMANGDIANTISPPYGLSPPPAAAEPMGYPODNFKQLDSY 259
Db 262 TTVNCIVEEVDARSVPYDEPLATGFVYMSPPGYR-EGSHTHTSYAADRFKQVDGF 320

Qy 260 FSDMLDRRKASLPVKRNFLLTSHFTVGVWDWAPKTRTVCSMTKWKVEYEMLRATVNGRYR 319
Db 321 YARDLTTKARATSPTRNLLTTPKFTVAVDWVPRPAVCTMTKQWEDMLRAEYGGSGFR 380

Qy 320 FMARELSATFISNTTEPDNRIILGOCIKREAAEAAIEQIFTKYNDSHVKVGHVQYFLAL 379
Db 381 FSSDAISTFTTNLTTEYSLSRVDLGDGICGRDAREIDRMFAKNATHYKVGQPYLYAT 440

Qy 380 GGFTVAYQPVLSKSLAHMYLRELNRDNTDMLDNVNNKHAIYKKNATLSRLRDRINA 439
Db 441 GGFLIAYQPLLSNTLAELYVREYMQ-----DRK-----PRNATP-----APLREA 482

Qy 440 PNRIITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSRIATAWELQNLRELVLNHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHQIRHVNDMLGRIVAWCELQNLHETLWNEA 539

Qy 500 IKINPSATASATLGRVVAAMGLDVAVSSCTAIDAESVTLONSMRVITTSNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMGLDGMVAVSTCVPAQDNVIVQNSMRVSSRPGTCYSRPLV 599

Qy 560 LFSYGENQGN-IOGOLGENNELPTLEAVEPCSAHRRYFLFGSGYALFENYFVKMVA 618
Db 600 SFRY-EDQGPLIEGOLGENNELRLTRDALEPCTVGHRRYFFGGYVYFEEYAYSHQLSR 658

Qy 619 ADIOIASTFVNLTLLEDEILPLSVYTKTELSDVGLDYAEVARNQLHELKFDYDINK 678
Db 659 ADVTTVSTFIDLNTMLDEHFEVPLEVYTRHEIKDGLDYTEVQRNQLHDLRFADIDT 718
```

Qy	679	VIEVDTN--YAGLOEF-----GCSMTKKDANPNNGOTQEAAARMEITDLI	722
		: :	
Db	719	VADANAAMFAGLCAFFEGMDLGRAVGKVMGVGGVVSAYSGVSSF-----MSNPf	772
		: :	
Qy	723	NAKAMTLASLDQYAKIEASLSAYSEATV-----NNLNLA---TLEQLKMAKTNLESAL	774
		: :	
Db	773	GALAVGLLVL-----AGLVAAPFAFYVLQIQRNPMKALYPUTTKELKTSDFPGVGGE	825
		: :	
Qy	775	NQANTDKTTIFONEHPNLVEA-----YKALKTTLE--QRATNLEGLSSTAYNQIRNNLVDL	827
		: :	
Db	826	GEEGAEGGGFD--EAKLAAREMIRYMALYSAMERTEHKARKKGTSALLSSKYVNMYLRK	883
Qy	828	YNKA 831	
Db	884	RNKA 887	

```

RESULT 8
US-10-237-551-18
; Sequence 18, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-237-551-18

```

	Query Match	30.98;	Score 1736;	DB 12;	Length 904;
	Best Local Similarity	43.08;	Pred. No. 1.5e-127;		
	Matches 363;	Conservative 142;	Mismatches 269;	Indels 70;	Gaps 16;
Qy	20	NSPSTQNVTSREVVSQVLSSEESTYLCPPVPVGVSTVIRLEPPRKCEPRKATEWGE	GI	79	
Db	82	NATVAAGHATRAHLRIKIKVENADAQYVCPPTGATVQFQPRRCPTPEQGNQYTE	GI	141	
Qy	80	AILFKENISPYKFKVLYLNIQTWTGTTRQITNRVTDTRDPVPSIEETLIDLGKR		139	
Db	142	AVFKEKNIAPYKFKATMYKDVIVSQVWFCHRSQFMGIFEDRAPVPEEVIDKIN	KV	201	
Qy	140	CSSKARYLRNNVVEAFDRDAGEKQVLLKPSKENTPESRAWHTTNETTYVMGSP	WYRTG	199	
Db	202	CRSTAKYVRNMETAFHRDDHETDMELKPAKVATRTSGWHITDLAYNSRVEAF	HYRG	261	
Qy	200	TSVNCVIEEMDARSVPYSYFAMANGDIANISPPYGLSPPEAAEPNGYPODN	FKQLDSY	259	
Db	262	TTVNCVIEEDARSVPYDFEVLATGDFVMSPPFYGYR-EGSHTETSYAADRE	KQVDGF	320	
Qy	260	FSDMLDKRRKASLPVKENFLTISHFTYGVDMWAPKTTVCSTWKKEYTEML	RAIVNGRYR	319	
Db	321	YANDLTTKATRTPTTNNLTTPKFTVADWVPKRPACVCTMKQWEYDEML	RAYGSGFR	380	
Qy	320	FMARELSATISNTEPDNRIILGQCICKREAAAEIQIFRTKYNDSHVKVQY	FLAL	379	
Db	381	FSSDAISTTTNLTYSLSRVDLGDICIGDAREADRMARKYNATHIKVQPO	YIYLAT	440	
Qy	380	GGFTVAYQPVLSKSLAHMYLRELMNRDRTDEMLDLVNKKHAIYKKNKTS	LSRLRDIRNA	439	
Db	441	GGFLIAYQPLLSNTLAEVLYREVMREQ-----DRK-----PRNATP	-----APLREA	482	
Qy	440	PNRKITLDDTTAKTSSTSSQVAMLFQYLDHIQHINDMFSRIATAWCELQ	NRELVLWHEG	499	
Db	483	PSANASVE---RIKTTSSIFARLQFTYNIQRHVDMLGRIAVAWCELQHT	LTWNEA	539	

```

QY 500 IKNPSTATSATLGRVAAKMLGDVAAVSSCTAIDAEVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAIASATVGRVRSARMMLGDVMAVSTCVVPADNVIVQNSMRVSSRPCTCYSRPLV 599
QY 560 LFSVGENQGN-IOGOLGENNELLTLEAVERPCSNHRRYFLEGSGYALPENYFNFKMVD 618
Db 600 SFRY-EDQGPLIEGOLGENNELRLTRDALEPCTVGHRRYFDFYFGGYYVFEEYAYSHQLSR 658
QY 619 ADIQIASTFVELNLTLEDREILPLSVYTTKEELRDVGVLDAEVARRNOLHELKEYDINK 678
Db 659 ADVTVTSTFDLNTMLTLEDHEFVPLEVYTRHEIKDGLDYTEVQRRNOLHDLREADIDT 718
QY 679 VIEYDTN---YAGLOEF-----GCMSTTKKDANPNNGOTQLEAARMELTDLI 722
Db 719 VIRADANAMFAGLCATFEFGMDLGRAVGKVMGVGVVSAVSGVSF-----MSNPF 772
QY 723 NAKMTLASLDYAKIEIASLSSAYSEATV-----NNNLNA-----TLEOLKMAKTNLESAI 774
Db 773 GALAVGLLVL-----AGLVAAPFAPRYVLQLQRNPKALYPLTTKELKTSDPGCVGGE 825
QY 775 NQANTDKTTEDNEHPNLEVA-----YKALKTTLE--QRATNLEGLUSTAYNQIRNNLVDL 827
Db 826 GEEGAEGGGFD--EAKLAAREMIRYMALVSAMERTEHKARKKGTFSALLSSKVTNMVLRK 883
QY 828 YNKA 831
Db 884 RNKA 887

RESULT 9
US-10-121-988-18
; Sequence 18, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITILE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRNT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

```

Query Match	30.98;	Score 1736;	DB 15;	Length 904;
Best Local Similarity	43.08;	Pred. No. 1.5e-127;		
Matches 363;	Conservative 142;	Mismatches 269;	Indels 70;	Gaps 16;
QY	20	NSSPSTQNVTSRREVSSVOLSEESTPYLCPPPPVGVSTVIRLEPPRKCPEPRKATWEGE	79	
Db	82	NATVAAGHATLRAHLREIKVENADAQVYVCPPTGATVQVFEQPRRCPTPEQONTYGEI	141	
QY	80	AILFKENISPYKFKVLYLYKNIQTWTGTYQTINRYTDRTPVSIEITLIDLGKGR	139	
Db	142	AVVFKENIAPYKFAWYKDVTVSQWFGHRSYQFMGIFEDRAPVPFEVIDKINTKGV	201	
QY	140	CSKARYLRNNVVEAFDRDAGEQVLLKSKKNTPE\$RAHHNTTETTYVWGSPIYRTG	199	
Db	202	CR\$TAKVVRNMET\$AFHRDHD\$TMDL\$KPAKVATRT\$RGWH\$T\$D\$KYNPSRVEAFHRYG	261	
QY	200	TSVNCIYEEMDARSVPFY\$YFAMANGDIANISPYGLSPPEAAEPMGYQQDNFKQIDSY	259	

[illegible]

QY 248 YPDNFKOLDYSFMDLDRKASLPVKRNELITSHFTVGVWDWAPKTRVCSMTWKVET 307
Db 284 EXPWFVSVNNKVVYDQNGRTVPLGKTRIFLDREYITLSWEKHUKLKNYSCLPLTLWKAFY 343
QY 308 EMRLATVNGRYRFAARELSATFIS-----NTTEFDPNRIILQCCKRAEAAIEQIF 359
Db 344 NGIQTEHSGSYHFVANDITASETTSKDKMKEFTTY-----HCLNEELKABIEKKY 394
QY 360 RTKYNDSHVGVHGYFIALGDFIVAYQPVLSKSLAHMYLRELNRDRTDEMLDVNNKH 419
Db 395 -AKVNSTHSKYGLKFKYKTDGGLYLWQPLIQ-----NRLLDKAKNLN 436
QY 420 AIYKKNATSLSLRDIRNAPNRKITLDDTTA---IKSTSSVQFAMQLFDYHDHIND 476
Db 437 -----NEYSRKRSQAESTDPMMEGTNGAGGEYSSENSTVAVQYAYDNLRIINN 491
QY 477 MFSRIATWCELQNLRELWHEGINKINPSATASATLGRVAAKGLDVAASVSCFAIDAE 536
Db 492 ILEDLSKAWCRQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIIVDQT 551
QY 537 SVTLQNSMRVIT-SYNTCYSRPLVLFSGYNOGNTQOGELNELLPTLEAVEPCSANHR 595
Db 552 SVSLHKLRLSLASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLATTVLETQENTE 611
QY 596 RYELFGSGYALFENYFNKVMYDAADIQIATSFVELNLLILEDEILPLSVYKELRDVG 655
Db 612 YFQAKTOMIYKNTYHEHLKTVPLSPSITLDTFIALNFTLLENVDVFKVLEYLTRDEKRLSN 671
QY 656 VLDYAEVARRNQLHKLKYDINK-VIEVDYTN-----AGLQBFQCMSTIKKDA--- 702
Db 672 VDIETMREYNYAQRVYSGRLKDLDLSTNRNQFVDAFGSLMDLGAQVQTVYVNAVSGV 731
QY 703 -----NPNNGQTOLEARMELTDLINAKMTLASLODYAKI--EASLSS 744
Db 732 ATLFSSIVTGFINKPFGG-----MLMIIWVIGVLFAYELTKKIYETAPIKM 783
QY 745 AYSEAEATVNNLN---ATLEQLKAKTNLESAINQANTDKTTFDNEHPNLVEAVKAKLT 800
Db 784 IYPEIDKLEKEREKEIAPISEEEELERIVLAHIIHQONSHMETKTRKOP-----KD 834
QY 801 TLEQRATNLEGLSSYAYNOIR-NLVDLYN 829
Db 835 SILTRAQNMLKRRS-GYSLNKAESVEMLN 863

RESULT 13

US-10-055-364-55
; Sequence 55, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:

; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/612,204

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US/60/142,736

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: US/60/168,532

; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 55

; LENGTH: 793

; TYPE: PRP

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Portion of Acelaphine herpesvirus.

US-10-055-364-55

Query Match 13.3%; Score 750; DB 14; Length 793;
Best Local Similarity 28.2%; Pred. No. 5.1e-50;

Matches 229; Conservative 131; Mismatches 362; Indels 90; Gaps 23;
QY 46 FYLCPPPVGSGVIRLEPPRCKPEPRKATATEWEGIAILEKENISPYKFKVLYYKNIIT 105
Db 12 FVCSASNIGDIFRQTSCHSCPNTKKEHNEGILLIEKENIVPVVFKVRKVIYVIST 70
QY 106 TWTGTTYQIYINRYDTRTPVSIIEITDLDIGKGRSSKARYLRNNVYVEAFDRDAGEKOV 165
Db 71 IYNGIYADAVTNQHFVSKSVPIYE-TRRMDITYOCYNSLDVTGNGLLVYTDNCGSMTV 129
QY 166 LLKPSKENTPESRAWHHTTNEYTVWGSPI---YRTGTSVNCIIEEMDARSVPFYSYAM 222
Db 130 DLQPVGDSNSVRRTHSQPEIHAEPG--WLLGGYRRRTVNCVETETDARAVPPRYFIT 187
QY 223 ANGDIANISPFYGLSPPEAAAEPMGYQDNFKQLDSYFMDLDRKASLPVKRNFLITS 282
Db 188 NIGDTIEMSPWKAWE--TEFSGEPDRTLTVAKDYRVVDYKFGTQPOGHTRFVDRKE 245
QY 283 HFTVGWDWAPKTRVCSMTKWKVEYTEMLEATVNGRYRFAARELSATFISNTTEFDPNRI 342
Db 246 EYTLSSAQOFRNISYCRWAHWSFDNAIKTEHGSLSHFVANDITASFYTPNTQ---TREV 302
QY 343 LGO--CIRKEAFAAEQIFRTKYNDSHVGVHGYFIALGDFIVAYQPVLSKSLAHMYL 400
Db 303 LKGVHCLNNTIESELKSL-ARKVNDTHSPNGTAQYLLTNGGLLVWQPLVQOKLLD--- 357
QY 401 ELMRDNRTDEMLDLYNNKHAIYKKNATSLSLRDIRNAPNRKITLDDTTAIKSTSSVOF 460
Db 358 -----AKGLLDVAKK-----QQNTTTTTRSRQRRSSVSGIDDDVYTAEST-- 403
QY 461 AMLQPLYDHIQTHINDMFSRIATWCELQNLRELWHEGINKINPSATASATLGRVAAKM 520
Db 404 TQIQFAYDTLRAQINNVLEELSRACRQHRASLMWNLKINPTSVMSIYGRVSAKR 463
QY 521 LGDVAASVSSCAIDAEVTLQNSMRV--ITSTNTCYSRPLVLFSGYNOGNTQOGELN 578
Db 464 IGDVISVSHCVVVDSDVSLHRSRMVPODKTHECYSRPVTFKFINDSHLYKGOGLVNN 523
QY 579 ELLPTLEAVEPCSANHRRYFLFGSGYALFENYFNKVMYDAADIQIATSFVELNLTLEDR 638
Db 524 EILLTTTAVTEICHENTEHYFQGGNNMYFKYRHVKTMPVGDVATLDTFVNLTLVNI 583
QY 639 EILPLSVYTKBELRDVGVDYAEVAR-----RNQLHEL-----KFYD----- 675
Db 584 DFQVIELYSREKRMSTAFDIETMREYNYTQRTVGLRLDRLDRLATNRNQFVDAFGSLM 643
QY 676 -----INKVIEVDYTAGLQFEGCMSTTKKDANPNNGQT--OLEAARMELTDLIN 723
Db 644 DDLGVVGKTVLNAVSSVATLFSSIVS-GIINFIK---NPFQGMLLFGLIAAVITVILN 699
QY 724 AKAMTLASLODYAKIEASLSAYSEAE--TVNNNLNATLEQLKAKTNLESA----- 773
Db 700 RKAKRFA--QNPVQMPIDIKTITSQRELOVDPIKSHELDRLMLAMHDYHASKQPEKQD 757
QY 774 INQANTDKTTFD-NEHPNLVE---AYKALKIT 801
Db 758 EEQGSTTSGPADLNKAKNVLRRRAGYKPLKRT 789

RESULT 14

US-10-055-364-40

; Sequence 40, Application US/10055364

; Publication No. US20020155433A1

; GENERAL INFORMATION:

; APPLICANT: Patience, Clive

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

; FILE REFERENCE: 61750-379

; CURRENT APPLICATION NUMBER: US/10/055,364

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/612,204

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US/60/142,736

; PRIOR FILING DATE: 1999-07-08

```

; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 40
; LENGTH: 844
; TYPE: PRN
; ORGANISM: Bovine herpesvirus 4
; US-10-055-364-40

```

Query Match 13.1%; Score 738.5; DB 14; Length 844;
Best Local Similarity 26.6%; Pred. NO. 4.6e-49;
Matches 217; Conservative 126; Mismatches 337; Indels 137; Gaps 22;

QY	19	TNSPSTQNV-TREVYVSSVOLSEESTFY---LCPPPVGS-TVIRLEPPRKCPEPRKAT	73
Db	62	STKSPSTNGOTSTPTPTTDDTASKNFYKRVCSASSSGELFRFLDLOTCPD-TKDK	120
QY	74	EMEGETAILFKENISPVKKVLYYKNIQTTWTGTTYRQITNRYTRDTPVSTEETIDL	133
Db	121	KHVEGILLVLKKNLPIFYFKRYKRTATSVTVYRGWSQAAYNRDDISRAIPYNELIS-M	179
QY	134	IDGKRCSSKARYLRNNVYVEAFDRDAGEQVLLKPKSFNTPESRAWHTTNTEYVWG-S	192
Db	180	IDRTYHCFSAMATVINGILNTYIDRSENKSVLPQVAGLTENIRYFSQPLIYAEPGW	239
QY	193	PWLYRTGTSYNCIVEEMDARSVPYSPFAMANGDIANISPF-----VGLSPPEAAE	244
Db	240	PGIYRVTVCEVMDIYAKSVPEYTHFIALGDITIEISPFCHNNSOCTTGNSTSRDATK	299
QY	245	PMGYPODNFKQLDSYFMDLDRKASPLVKRNFLLITSHFTVGDWDAKPTRVCSMTKWK	304
Db	300	V-WIENHQTVDY-----ERRGHTPKDKRFLKDEEYTSWKAEDRERAICDFVIW	350
QY	305	EVTEMLRATVNGRYRFMARLSATFISNT---TEFDPNRILLGOCIKREABAEIQIFRT	361
Db	351	TFPRAIQTIHESFHFVANEVTSFSLTSNOEETELRGTEIL-NCMNSTINETLEETVK-	408
QY	362	KYNDSRVKGHVGYFLAAGFYVAYQVPLSKSLAHMYLRELMRDNRDTEMLDVNNKHAI	421
Db	409	KFNKSHIRDEGVYKYTNGGLFLIQWAMKPLNISEHTNYTIERNKTI-----GNK---	458
QY	422	YKNATSLSLRRDIRNAPNRKTILODPTAIKTSTSSVOFAMLOFLYDHIQTHINDMFRI	481
Db	459	-----SROKRSV-----DTKTFQAGKGLSTAQOVAYDHLRTSMNHILEEL	499
QY	482	ATAWCELQNELVLWHEGINKINPSATASATLGRVAAKMLGDVAVSSCTAIDAEVTLQ	541
Db	500	TKTWCREQKDNLMWYELSKINPSVMAAIIYGPVAVKMGDAFPWSECINVDQASVNIH	559
QY	542	NSMRVITSTWTCYSRPLVLFVSYGNOGNIQGLGENNELPTLEAVPSCSANHRRYFLFG	601
Db	560	KSMT--DDPKVCYSRPLVTFKVNSTATFRQLGTRNEILLTHTVETCRPTADHYFVK	618
QY	602	SGYALPENYFVKWDAADTOIASTFVELNLTLEDERIILPLSVYTKHEELRDGVGLDYAE	661
Db	619	NMTHYFKDYKFXVTMDNTNISTDLFTLLNFTIDNIDFKTVELYSETERKMASALDLET	678
QY	662	VARRNQLHELKEYDINKVIE--VDNTYAGLQERGCMSITKKDANPNNGQTLEARMELT	719
Db	679	MFREYNYTQKLASLRLEDLNTIDLN-----RDR-----	707
QY	720	DLINAKAMTLASLDQYAKIEBASLSYSEAEVY-----	752
Db	708	LVKDLSEMMADLGDIGKV--VVNTFSGIYTVFGSVGGVFSFPTNPIGVVTIILLIIV	763
QY	753	-----NNNLNATLEBOLKMAKNTLESAINQAN	778
Db	764	VFWFVIVTSRRTNMNEA--PIKMIYFNIDKASEQEN	798

RESULT 15
US-10-055-364-38

```

: Sequence 38, Application US/10055364
: Publication No. US20020155433A1
: GENERAL INFORMATION:
: APPLICANT: Patience, Clive
: TITLE OF INVENTION: Gamma Herpes V1
: FILE REFERENCE: 61750-379
: CURRENT APPLICATION NUMBER: US/10/0
: CURRENT FILING DATE: 2009-01-23
: PRIOR APPLICATION NUMBER: US/09/612
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US/60/142
: PRIOR FILING DATE: 1999-07-08
: PRIOR APPLICATION NUMBER: US/60/168
: PRIOR FILING DATE: 1999-12-02
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 38
: LENGTH: 808
: TYPE: PRT
: ORGANISM: Rhesus monkey rhadinovir
US-10-0355-364-38

```

Query Match 12.9%; Score 725.5; DB 14; Length 808;
Best Local Similarity 29.2%; Pred. No. 4.5e-48;
Matches 208; Conservative 115; Mismatches 306; Indels 83; Gaps 19;

```

Qy 5 RRCIFFLIVILYGTNSSPSTQNTVS-REVVSSVQLSEEST-----FYLC 49
Db 8 RLLLRWVVIIAIGTAVG---ENVTTKPGATTAKPTPGPSTPTPENPRAEAFKFRVC 64
Qy 50 PPPVGSVIRLEPPRKCEPRKATEWEGEGTALIFKENISPKRKVTIYKNIQTITWTG 109
Db 65 SASATGELFRNLEKTCPGTEGKTH-OEGILMFWKKNIVPHIFKVRRYKVAITSVTYRG 123
Qy 110 TTYRQIINRYTDRTPVSEIETLIDIDGKSCSKARYLRNNVVEAPDRDAGERQVLLKP 169
Db 124 WTEATVTKQEVIRPVPQYEINH-M-DTYOCFSSMRVNVANGIVNTYTRDFTNQTVLQP 182
Qy 170 SKFNTPESRAWHHTNETHYVWG-SPIYIRGTSTVNCIVEEMDARSPEPPSYFAMANGDIA 228
Db 183 VEGILTNDIQRYFSQPVLYTTPGFWPGGIYRVYRTVNCBIVDMIARSABSPYFVFTALGDTV 242
Qy 229 NISPF----XGLSPPEAAAPMGYPQDNFKQLODSYFSMDLDRKRKASLPVKRNFLLTSHF 284
Db 243 EVSPFCHNDSTCSVAEKTENGLG-----ARVLNNTYMDVAFR--APTTRTRVFADSGEY 295
Qy 285 TVGMDWAPKTRVCSMTKWEKTEMLRATVNGRYRFRMAELSATF---ISNTTFEFDNRI 341
Db 296 TVSKAEDPKSANCALTLWNTFPRAIQTTHIASYHFVANDYTAFTFSPLESEVANFTCTYS 355
Qy 342 ILGCQIKREAEAEIQFRTKYNDSHVKGVHGVQYFALGGFIYAYQVPLSKSLAHMYLRE 401
Db 356 CLDEVIQKTLNDTIK-----KLSDTHVTNGSAQVYKTEGGFLLELLWQLPLSL----- 403
Qy 402 LMRDNRFTDMDLVLNNKHAIYKKNATSLRLRDIRNAPNRKTTLDOTTAIKSTSSVOFA 461
Db 404 -----VDEMRLG-NGTTPAPPATTSTVSRRRSV-----NTNEQATDNLAAAP 444
Qy 462 MLOFLYDHIQTHINDMFSTRATACELQNELVLWHEGIKINPSATASATLGRVAAKML 521
Db 445 QLOFAYDKLRASINKVLEELSRWCRCQVRDXTYMYELSKINPTSVMTAIYGRPVSAKFV 504
Qy 522 GDVAAVSSCTAIDAESVTLQNSRVITST-NTCYSRPLVLFSGYGENOGNQGOLGENNEL 580
Db 505 GDAISVTDCAVDAQASVSIHSLR--TSTFGMCSYRPPVTFRFLNSITLTFEGQLGPRNEI 562
Qy 581 LPTLEAVEPCSAHRRYFLFGSGYALFENFYKRVDAADIQIASTFVELNLTLEDDREI 640
Db 563 ILTDNQVEACKETCEHYFIASNTYYKYDVFVKKINTSEISTGLTFIALNLSFENIDF 622
Qy 641 LPLSVYTKKELRDVG-VLDYAEVARRNQLHELKPYDINKVIEDVNTYVAGLOE 691
Db 623 RVLEIYSRAEKKISGVSFDIETMFR-----EYNY-----TORLAGURE 661

```

RESULT 15
US-10-055-364-38

Search completed: October 8, 2003, 17:28:07
Job time : 65.4796 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:40 ; Search time 30.5244 Seconds
(without alignments)
1505.341 Million cell updates/sec

Title: US-09-147-052-4
Perfect score: 5619
Sequence: 1 MHYFRNCIFFLIVILYGTN.....SSNENADKIPYRPGTFL 1086

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3601	64.1	865	1	US-07-803-633A-13
2	1968	35.0	879	1	US-08-220-151-2
3	1968	35.0	879	1	US-08-220-151-3
4	1968	35.0	879	1	US-08-413-118-2
5	1968	35.0	879	1	US-08-413-118-3
6	1968	35.0	879	1	US-08-413-118-106
7	1968	35.0	879	3	US-08-473-446-2
8	1968	35.0	879	3	US-08-473-446-3
9	1968	35.0	879	3	US-08-473-446-106
10	1913	34.0	615	2	US-08-525-742-10
11	1859	33.1	610	2	US-08-525-742-8
12	1846	32.9	1041	1	US-08-220-151-4
13	1846	32.9	1041	1	US-08-413-118-4
14	1846	32.9	1041	3	US-08-473-446-4
15	1779	31.7	913	3	US-09-232-468A-2
16	1779	31.7	913	4	US-09-784-984B-49
17	1775.5	31.6	868	1	US-08-220-151-7
18	1775.5	31.6	868	3	US-08-413-118-7
19	1775.5	31.6	868	3	US-08-804-439A-21
20	1775.5	31.6	868	3	US-08-473-446-7
21	1775.5	31.6	868	3	US-08-720-229-21
22	1775	31.6	913	1	US-08-220-151-6
23	1775	31.6	913	1	US-08-413-118-6
24	1775	31.6	913	3	US-08-473-446-6
25	1755	31.2	913	6	5196516-8
26	1748	31.1	933	2	US-08-682-847-2
27	1743	31.0	943	3	US-08-911-321-4

28 1741.5 31.0 980 1 US-08-220-151-5 Sequence 5, Appl1
29 1741.5 31.0 980 1 US-08-413-118-5 Sequence 5, Appl1
30 1741.5 31.0 980 3 US-08-473-446-5 Sequence 5, Appl1
31 1739.5 31.0 891 1 US-08-042-747A-6 Sequence 6, Appl1
32 1736 30.9 904 4 US-09-894-998A-18 Sequence 18, Appl1
33 1735 30.9 904 3 US-08-632-537-2 Sequence 2, Appl1
34 1735 30.9 904 5 PCT-US96-05316-2 Sequence 2, Appl1
35 1735 30.9 904 6 5244792-3 Patent No. 5244792
36 1719.5 30.6 885 1 US-08-042-747A-8 Sequence 8, Appl1
37 1714.5 30.5 885 3 US-08-804-439A-23 Sequence 23, Appl1
38 1714.5 30.5 885 3 US-08-720-229-23 Sequence 23, Appl1
39 1705.5 30.4 904 3 US-08-632-537-1 Sequence 1, Appl1
40 1705.5 30.4 904 5 PCT-US96-05316-1 Sequence 1, Appl1
41 1702.5 30.3 903 3 US-08-804-439A-22 Sequence 22, Appl1
42 1702.5 30.3 903 3 US-08-720-229-22 Sequence 22, Appl1
43 1699.5 30.2 903 1 US-08-220-151-8 Sequence 8, Appl1
44 1699.5 30.2 903 1 US-08-413-118-8 Sequence 8, Appl1
45 1699.5 30.2 903 3 US-08-473-446-8 Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025ofu
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, YI
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-803-633A-13

Query Match 64.1%; Score 3601; DB 1; Length 865;
Best Local Similarity 85.9%; Pred. No. 1.5e-236;
Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;

QY

1 MHYFRNCIFFLIVILYGTNVSPTONTVSREVSVQLSEESTFYLCPPPVGSTVRL 60


```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: US/08/413.118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220.151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
; US-08-413-118-2

Query Match      35.0%; Score 1968; DB 1; Length 879;
Best Local Similarity 45.8%; Pred. No. 1.7e-125;
Matches 387; Conservative 153; Mismatches 253; Indels 52; Gaps 9;

QY 21 SSPSTQNTVTSREWSSVOL-SEESTFYLCPPVGVSTVIRLEPPKRCPEPRKATWEGGI 79
DB 36 STPKPTSDIREILRESQIESDDTFTYMCPPSGSTLVRLEPPRACPNYKLGKNTFEGI 95
QY 80 AILFKENTSPKFKVLYKNIQTITWTGTYRITNRYTDRTVPVSEETLIDLIDGKR 139
DB 96 AVIFKENTSPKFKANIYKNIITVWSSGTAVITNRYTDRTVPVSEETLIDLIDRGM 155
QY 140 CSSKARYLNRYVYFAFDKDEPQVLLKPKFNTPESSAWHTTNETYTVMGSPWYRTG 199
DB 156 CLSKADYRNRYVYFAFDKDEPQVLLKPKFNTPESSAWHTTNETYTVMGSPWYRTG 215
QY 200 TSVNCIVEEMDARSVPYFAMANGDIANTSPFYGLSPPEAAEPMGYPODNFKOLDSY 259
DB 216 TSVNCIVEEMDARSVPYFAMANGDIANTSPFYGLSPPEAAEPMGYPODNFKOLDSY 274
QY 260 FSWDLKRRKASLPVKNRFLITSHFTVGVNDWAPKTRVCSMTKKEVTEMLRATVNGRYR 319
DB 275 YFIDLDLRLQLGAPVSRNFLTQHTVYVANNVWPKIREVCTLAKWEIDEILRDEYKGSYR 334
QY 320 FWARELSATFISNTTEFDNRIILQCCKKREAAEIQFRTKYNDSHVKVGHVYFLAL 379
DB 335 FTAKSISATFISDTTQFDIDRVKLSDCAKREAEIAIDKIYKKYKNTHTIQTGETYLAR 394
QY 380 GGFIVAYOVLKSLAHYVLEMLNDRTDMLDVLNKNHAIYKKNATSLSLRLRDINA 439
DB 395 GGFIAFRPMISNELAKLINELVNSRNTVDUKSLNFSVRGAKRRRSVEENKSKRNI 454
QY 440 PNRKITDLDTTAISKSSYQFAMLOFLYDHIOTHIINDMFSTRATACELQNRRELVWHEG 499
DB 455 EGGIENVNNTTIKTTSSVHFAMLOFAYDHIQSHVNMELSRATACWNLQNKERTLWNEV 514
QY 500 IKINFSATASVILGRVAAKMLGDVAVSSCTAIDAESVTGLONSRRVITSTWTCYSRPLV 559
DB 515 MFLNPTSVASVAMDQVRSARMLGDVLAVTQCYNISGSSVFIQNSRVLGSLTTCYSRPLI 574
QY 560 LFSYGNOCN-IQGLGENNELLPLEAVEPCSAHRRYFLFGSGVAFENYFNKMWDA 618
DB 575 SFKALENTNYLEGQGENNELLVKLEPCTAHKRYFKEGADYVFENYAYVKVPL 634
QY 619 ADIQIASTFVELNLTLEDRLPLSVYTKELRDVGVDLYDAEVARRNQLHKLKPYDINK 678

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: US/08/413.118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220.151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-3

Query Match      35.0%; Score 1968; DB 1; Length 879;
Best Local Similarity 45.8%; Pred. No. 1.7e-125;
Matches 387; Conservative 153; Mismatches 253; Indels 52; Gaps 9;

QY 21 SSPSTQNTVTSREWSSVOL-SEESTFYLCPPVGVSTVIRLEPPKRCPEPRKATWEGGI 79
DB 36 STPKPTSDIREILRESQIESDDTFTYMCPPSGSTLVRLEPPRACPNYKLGKNTFEGI 95
QY 80 AILFKENTSPKFKVLYKNIQTITWTGTYRITNRYTDRTVPVSEETLIDLIDGKR 139
DB 96 AVIFKENTSPKFKANIYKNIITVWSSGTAVITNRYTDRTVPVSEETLIDLIDRGM 155
QY 140 CSSKARYLNRYVYFAFDKDEPQVLLKPKFNTPESSAWHTTNETYTVMGSPWYRTG 199
DB 156 CLSKADYRNRYVYFAFDKDEPQVLLKPKFNTPESSAWHTTNETYTVMGSPWYRTG 215
QY 200 TSVNCIVEEMDARSVPYFAMANGDIANTSPFYGLSPPEAAEPMGYPODNFKOLDSY 259
DB 216 TSVNCIVEEMDARSVPYFAMANGDIANTSPFYGLSPPEAAEPMGYPODNFKOLDSY 274
QY 260 FSWDLKRRKASLPVKNRFLITSHFTVGVNDWAPKTRVCSMTKKEVTEMLRATVNGRYR 319
DB 275 YFIDLDLRLQLGAPVSRNFLTQHTVYVANNVWPKIREVCTLAKWEIDEILRDEYKGSYR 334
QY 320 FWARELSATFISNTTEFDNRIILQCCKKREAAEIQFRTKYNDSHVKVGHVYFLAL 379
DB 335 FTAKSISATFISDTTQFDIDRVKLSDCAKREAEIAIDKIYKKYKNTHTIQTGETYLAR 394
QY 380 GGFIVAYOVLKSLAHYVLEMLNDRTDMLDVLNKNHAIYKKNATSLSLRLRDINA 439
DB 395 GGFIAFRPMISNELAKLINELVNSRNTVDUKSLNFSVRGAKRRRSVEENKSKRNI 454
QY 440 PNRKITDLDTTAISKSSYQFAMLOFLYDHIOTHIINDMFSTRATACELQNRRELVWHEG 499
DB 455 EGGIENVNNTTIKTTSSVHFAMLOFAYDHIQSHVNMELSRATACWNLQNKERTLWNEV 514
QY 500 IKINFSATASVILGRVAAKMLGDVAVSSCTAIDAESVTGLONSRRVITSTWTCYSRPLV 559
DB 515 MFLNPTSVASVAMDQVRSARMLGDVLAVTQCYNISGSSVFIQNSRVLGSLTTCYSRPLI 574
QY 560 LFSYGNOCN-IQGLGENNELLPLEAVEPCSAHRRYFLFGSGVAFENYFNKMWDA 618
DB 575 SFKALENTNYLEGQGENNELLVKLEPCTAHKRYFKEGADYVFENYAYVKVPL 634
QY 619 ADIQIASTFVELNLTLEDRLPLSVYTKELRDVGVDLYDAEVARRNQLHKLKPYDINK 678
```


Db 575 SFKALNSTNYIEGOLGENNELVERKLEPCTANHKRYFKGADYVYFENYVYKVP 634
 QY 619 ADIQIATFVELNLTLEDREILPLSVYTKELRDVGVDYAEVARRNOLHELKFDINK 678
 Db 635 NEIEMISAYVDLNTLLEDREFLEVTYTRAELEDTGLDYSEIQORNLHAKFYDIDS 694
 QY 679 VIEVDNTAGLOEFGCMSITTKDANPNNGOTOLEARMELTDLINAKAMTLASLODYAK- 737
 Db 695 VVKVDNNVIMRGIANFQGLGDVGAGKGVVLGAA-----NAVATVSGVSSFLNN 746
 QY 738 -----IBASLSAYSEAEVTNNNLNATLEOLKMAKT-NLESAINQANTDKTTFDN 786
 Db 747 PFGALAVGLLIAGLFAAFAYRYVSKLSNPKALYPVTTKNLAKESVKNNGSNGSDGE 806
 QY 787 EHPNLVE-----AYKALKTTLEOR-----ATNLEGLS-----STAY 817
 Db 807 ENDDNIDEKLOQAKEMIKYMSLVAMEQOEHAIKKNSGPGALLASHITNLSLKHGPKY 866
 QY 818 NOIRN 822
 Db 867 KRLKN 871

RESULT 7

US-08-473-446-2
 ; Sequence 2, Application US/08473446
 ; Patent No. 6017542
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/473,446
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/413,118
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FROMMER, WILLIAM S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2670
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 879 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal

US-08-473-446-2
 Query Match 35.0%; Score 1968; DB 3; Length 879;
 Best Local Similarity 45.8%; Pred. No. 1.7e-125;
 Matches 387; Conservative 153; Mismatches 253; Indels 52; Gaps 9;

QY 21 SSPSTQNTVTSRVVSSVOL-SEESTFYLCPPPVGSTVIRLEPPKCPPEPRKATEWGEIGI 79
 Db 36 STPKPTSDDIRILRESQIESDDTSTFYMCPPSGSTVLRLEPPRACPNYKLGKNFTEGI 95
 QY 80 ALLKENTISPYKFKVLYKNIQIOTWTGTYRTITRYTDTPTVSIIEITDLIDGKGR 139
 Db 96 AVIFENISPYKFKANIKYKNIITVWSGSYIVITNRYTDRVPIGVFEITELIDRRM 155
 QY 140 CSSKARYLRNNVYVFAFDRDAGEKQVLLKPSKENTPESRAWHTTNTETVWGSWPYTRTG 199
 Db 156 CLSKADYIRNNYEFTAFDKDEDPREVHLKPSKENTPGSRGHTVNDYTKIGSGSYHSG 215
 QY 200 TSVNCIVEEMDARSVPYSPYFAMANGDIANISPPFVGLSPPEAAEPMPQPNFQOLDY 259
 Db 216 TSVNCIVEEDARSVPYSPYFAMANGDIANISPPFVGLSPPEAAEPMPQPNFQOLDY 274
 QY 260 FMSMDLDRKRSKASLPYKRNFLITSHFTVGMWDMPKTRVCSMTKWKEVTMLRATVNGRYR 319
 Db 275 YPIDLTLRLQGLGAPVSRNFLTTHQVTVANWVVPKIREVCTLAKWEIDELIIRDEYKGSYR 334
 QY 320 FMARELSATFISNTTFDPNRIILGQCIRKREAAAEQIFRTKYNDSHVKGHVQYFLAL 379
 Db 335 FTAKSISATFISDTTQFIDRVKLSDCAKREAEIAIDKIYKKYKNTHTIQTGELEYLAR 394
 QY 380 GGFIVAYQPVLSKSLAHMYLRELMDNRDMDLNVNKHAIYKKNATSLRLRDIRNA 439
 Db 395 GGFIIAFRPMISNELAKLYINELVRSNRTVDLKSLLNPSVRGBGAKRRKRSVEENKSRNI 454
 QY 440 PNKRITLDDTTAKTSSVOFAMLOFLYDHIOTINDMFSTRATACELQNLRELVLWHS 499
 Db 455 EGIENVNNTIKTSSVHFAMLOFAYDHIQSHVNMELSRATACNQLKERTLWNEV 514
 QY 500 IKINPSATASATLGRVAAKMLGVAVSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
 Db 515 MKLNPTSVASVANDQVRSARMLGDLAVTQCVNISGSSVFQNSMRVLSTTTCYSRPLI 574
 QY 560 LFSYENOGN-IQGGOLGENNELPLLEAVEPCSAHRRYFLPGSYALFENTNFVKWDA 618
 Db 575 SFKALNSTNYIEGOLGENNELVERKLEPCTANHKRYFKGADYVYFENYVYKVP 634
 QY 619 ADIQIATFVELNLTLEDREILPLSVYTKELRDVGVDYAEVARRNOLHELKFDINK 678
 Db 635 NEIEMISAYVDLNTLLEDREFLEVTYTRAELEDTGLDYSEIQORNLHAKFYDIDS 694
 QY 679 VIEVDNTAGLOEFGCMSITTKDANPNNGOTOLEARMELTDLINAKAMTLASLODYAK- 737
 Db 695 VVKVDNNVIMRGIANFQGLGDVGAGKGVVLGAA-----NAVATVSGVSSFLNN 746
 QY 738 -----IBASLSAYSEAEVTNNNLNATLEOLKMAKT-NLESAINQANTDKTTFDN 786
 Db 747 PFGALAVGLLIAGLFAAFAYRYVSKLSNPKALYPVTTKNLAKESVKNNGSNGSDGE 806
 QY 787 EHPNLVE-----AYKALKTTLEOR-----ATNLEGLS-----STAY 817
 Db 807 ENDDNIDEKLOQAKEMIKYMSLVAMEQOEHAIKKNSGPGALLASHITNLSLKHGPKY 866
 QY 818 NOIRN 822
 Db 867 KRLKN 871

RESULT 8

US-08-473-446-3
 ; Sequence 3, Application US/08473446
 ; Patent No. 6017542
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.

```

; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-473-446-3

Query Match 35.0%; Score 1968; DB 3; Length 879;
Best Local Similarity 45.8%; Pred. No. 1.7e-125;
Matches 387; Conservative 153; Mismatches 253; Indels 52; Gaps 9;

QY 21 SSPSTQNTVSREVSVSQL-SEESTFYLCPPVGVSTVIRLEPRKCPERKATWEGGI 79
DB 36 STPKPSTDDIRETLRESQESDDTSTFYMCPPPGSGTLVRLPRACPNYKLGKNETGI 95
QY 80 AILFKENISPKFKVLYKNIQTWTGTTTQRTNRYTDRTPVSEIEITLIDCKGR 139
DB 96 AVIFKENISPKFKANIYKNIITVWGSTVAVITNRYTDRVPGVPEITELIDRRGM 155
QY 140 CSSKARYLRNVVVEAFDRDAGEKQVLLKPSKENTPESRAWHHTTNETYTVWGSPIWYRTG 199
DB 156 CLSKADYIRNNYETFAEDKDEDPREVLHLPKSKENTPGSRGWHVNDYTYKIGSGGYHSG 215
QY 200 TSVNCIVEEMDARSVPYFYFAMANGDIANISPFYGLSPPEAAAPMGYPQDNFKQLDSY 259
DB 216 TSVNCIVEEDARSVPYDYSFALSTGDIHMSFFGLR-DGAHTEVISYSTDRFQIEGY 274
QY 260 FSDLDKRRKASLPVRNFLTSHFTVGDWDAPKTTTRVCSMTKWEVTEMLRATVNGRYR 319
DB 275 YPIDLTRLQLGAPVRNFLTTHQVTVANNVWPKIREVCTLAKWREIDEIRDEYKGSYR 334
QY 320 FMARELSATPISNTTFDPNRILGQIKREAAAEQIFRTKYNDSHVKGHVQYFLAL 379
DB 335 FTAKSISATFISDTTFQDIDRVKLSDCAKREATEADKYKKYKNTHTIQTGLELYLAR 394
QY 380 GGFIVAYQPVLSKLAAHMYLRELNRDNRDMDLVLNNKHAIYKKNATSLSLRLDRIRNA 439
DB 395 GGFIIAFRPMISNELAKLYINELVRSNRTVDLKSLLNPSVRGCARKRRSRSEENKSRKNI 454
QY 440 PNKRIILDDTTAKTSSTSSVQFAMQLQFLYDHIQHINDMFSRIATANCELQNLVWHEG 499
DB 455 EGGIENVNNTSIKTTSSVHFAMQLQFAYDHIQSHVNEMLSRATATCNLQNKERTLWNEV 514
QY 500 IKINPSATASATLGRVAAKMLGDVAASVSCATAIDAEVTLQNSMRVITSTNTCYSRPLV 559
DB 515 MKNLNPYSVASVAMDQRVSAKMLGDVLAVTQCNVISGSSVFIQNSMRVLGSTTTTCYSRPLI 574

```

```

QY 560 LFSYGENQGN-IQGQGENNELPTLEAVEPCSANHRRYFLFGSGYALPENYFVXWDA 618
DB 575 SFKALENSTNYIEGQGENNELVERKLIPECTANHKRYFKFGADYVYFENYAVRKPVL 634
QY 619 ADIOIASTFVELNLTLLEDEIREILPLSVYTKKEERDVGVLDAEVARRNOLHELKFFVDINK 678
DB 635 NEIEMISAVVDLNLTLLEDEIREFLPLEVYTRAELEDYGLDYSEIQRNQJHALKFYDIDS 694
QY 679 VIEVDNYAGLQFQCMSTTKKDANPNNGQTOLEAAARMELTDLINAKAMTLASIQDYAK - 737
DB 695 VVKYDNNVIMRGIANFFQGLGDVGAGFGKVLGAA-----NAVIAVSGVSSFLNN 746
QY 738 -----IEASLSAYSEATVNNNLNATIEQLKMAKT-NLESAINQANTDKTTFDN 786
DB 747 PFGALAVGLLILAGLFAAFLAYRYVSKLSNPKALYPVTTKNLKESVKNGSGNNSDGE 806
QY 787 EHPENLVE-----AYKALKTTLEQR-----ATNLEGLS-----STAY 817
DB 807 ENDNIDEKLOAQKEMIKYNSLVSAEQOEHKAIKKNSGPALLASHITNLSLKHGPKY 866
QY 818 NQIRN 822
DB 867 KRLKN 871

```

RESULT 9

```

US-08-473-446-106
; Sequence 106, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PROLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-473-446-106

```

Query Match

35.0%; Score 1968; DB 3; Length 879;

[illegible]

RESULT 10

RESOLV 10
US-08-525-742-10

05 00 215 742 10
; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

APPLICANT: Saito, Shuji

APPLICANT: Ohkawa, Setsuko

APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

```

1  APPLICANT: Funato, Hirono
2  APPLICANT: Iritani, Yoshikazu
3  APPLICANT: Aoyama, Shigemichi
4  APPLICANT: Takahashi, Kiyoochito
5  TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
6  TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BE
7  TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RE
8  TITLE OF INVENTION: AS USE THEREOF
9  NUMBER OF SEQUENCES: 51
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
12 ADDRESSEE: NAUGHTON
13 STREET: 1725 K Street, Suite 1000
14 CITY: Washington
15 STATE: DC
16 COUNTRY: USA
17 ZIP: 20006
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/525,742
25 FILING DATE: 25-SEP-1995
26 CLASSIFICATION: 424
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: JP 05-074139
29 FILING DATE: 31-MAR-1993
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: JP 05-245625
32 FILING DATE: 30-SEP-1993
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: PCT/JP94/00541
35 FILING DATE: 31-MAR-1994
36 ATTORNEY/AGENT INFORMATION:
37 NAME: McLeLland, Le-Nhung
38 REGISTRATION NUMBER: 31,541
39 REFERENCE/DOCKET NUMBER: 950811
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 202-659-2930
42 TELEFAX: 202-8870357
43 INFORMATION FOR SEQ ID NO: 10:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 615 amino acids
46 TYPE: amino acid
47 TOPOLOGY: linear
48 MOLECULE TYPE: protein
49 US-08-525-742-10

```

Query Match

Best Local Similarity 95.9%; Pred. No. 5.6e-122;

Best local similarity	90.78;	Freq.: NO; 3.0E-122;
Matches	377;	Conservative
Matches	6;	Mismatches
Indels	0;	Gaps

QY 684 TNYAGLQEFFGCMSTTKKDANPNNGQTQLEAARMELTDLINAKAMTLASLQDYAKIEASLS 743

[illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

744 SAYSEAETVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803 QY

Db 77 SAYSEAE TVNNLNATLEQLKMAKTNLESA INQANTDKTTFDNEHPNLVEAYKALKTTLE 136

100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 137 QRATNLEGLASTAYNQIRNNLVLDLYNNASSLITKTLDPNGGMLDSDNEITTVNRNINNT 196

QY 864 LSTINEQKTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKY 923

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 984 AADANNVGLQYKLNNGNVQVFEATSTANNTTANPTPAVDKIVAKIVLSGLRFGONTI 1043
 DB 317 AADANNVGLQYKLNNGNVQVFEATSTANNTTANPTPAVDKIVAKIVLSGLRFGONTI 376
 QY 1044 ELSVPTGEGNMKNKVPAMIGNIYLSNENNAKDI 1076
 DB 377 ELSVPTGEGNMKNKVPAMIGNIYLSNENNAKDI 409

RESULT 11
 US-08-525-742-8
 ; Sequence 8, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hiroo
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemitsu
 ; APPLICANT: Takahashi, Kiyochito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSER: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742
 ; FILING DATE: 25-SEP-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-074139
 ; FILING DATE: 31-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-245625
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP94/00541
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mclelland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 950811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-659-2930
 ; TELEFAX: 202-8870357
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 610 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-525-742-8

Query Match 33.1%; Score 1859; DB 2; Length 610;
 Best Local Similarity 93.4%; Pred. No. 2.6e-118;
 Matches 366; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 684 TNYAGLOEPCGCHSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLODYAKIEASLS 743
 DB 17 TSFLSIGISSCHSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLODYAKIEASLS 76
 QY 744 SAYSEATVNNLNATLEOKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803
 DB 77 SAYSEATVNNLNATLEOKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 136
 QY 804 QRATNLEGLSTAYNOIRNLDVLYNKASSLITKTLPLNGGTLLDSNEITTANKNNNT 863
 DB 137 QRATNLEGLSTAYNOIRNLDVLYNKASSLITKTLPLNGGTLLDSNEITTANKNNNT 196
 QY 864 LSTINEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKY 923
 DB 197 LSTINEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKY 256
 QY 924 ARRTVMNGDEPSRILANTNSITDVSIIYSLAGTNTKYQSFNSYNSGPGSTGYLYFFPYKLVK 983
 DB 257 ARRTVMNGDEPSRILANTNSITDVSIIYSLAGTNTKYQSFNSYNSGPGSTGYLYFFPYKLVK 316
 QY 984 AADANNVGLQYKLNNGNVQVFEATSTANNTTANPTPAVDKIVAKIVLSGLRFGONTI 1043
 DB 317 AADANNVGLQYKLNNGNVQVFEATSTANNTTANPTPAVDKIVAKIVLSGLRFGONTI 376
 QY 1044 ELSVPTGEGNMKNKVPAMIGNIYLSNENNAKDI 1075
 DB 377 ELSVPTGEGNMKNKVPAMIGNIYLSNENNAKDI 408

RESULT 12
 US-08-220-151-4
 ; Sequence 4, Application US/08220151
 ; Patent No. 5529780
 ; GENERAL INFORMATION:
 ; APPLICANT: Paoletti, Enzo
 ; APPLICANT: Limbach, Keith J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc AND gd AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 91
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Curtis, Morris & Safford
 ; STREET: 530 Fifth Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/220,151
 ; FILING DATE: 30-MAR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Frommer, William S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2540
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1041 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: N-terminal
 ; US-08-220-151-4

```

Query Match      32.9%: Score 1846; DB 1; Length 1041;
Best Local Similarity 40.4%: Pred. No. 4.4e-117;
Matches 384; Conservative 158; Mismatches 259; Indels 150; Gaps 13;

Qy 18 GTNPSPTQNTVSREVSSVOLSEEE-----STFYLCPPVPVSTVIRLEPPRKCPEPRKAT 73
Db 93 GTPPKPTDPTDMSDREALRASQIEANGPSTFYMCPPSPGVVLEPPRACPDYKLGK 152
Qy 74 EWGEGIALPKENISPKFKVLYKNIQTITWTGTYRQITNRYTDRTPVSEETDL 133
Db 153 NFTGIAVIFKENIAPYKFKANIYKNIIMTVWSSGSYAVTTRNRYTDRVPVKVQETDL 212
Qy 134 IDGRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESR----- 178
Db 213 IDRGMCLSKADYVYRNNYQFTAFDRDEDPRELPLKPSKFNTPQSRGHWYKFKATVYK 272
Qy 179 ----- 178
Db 273 VIVSTANAGSSYTOITNRYADRPVPIVSEITDITDKFGKSSKATYVYRNNHKVFAFNEDK 332
Qy 179 -----AMHTTNETYVWGSPIYRTGTSVNCEIVEEDARSVPYYSYF 220
Db 333 NPQDMLIASKYNSVGSKAHHTTNETYKGAAGFHHSQTSVNCIVEEDARSVPYDSF 392
Qy 221 AMANGDIANISPFYGLSPPEAAEPGYPQDNFQDLSYFSMDLDRKRSKASLPVKRNF 280
Db 393 AISTGDVHNSPFFGLR-DGAHVHTSYSSDRFQOIEGYPIDLDTLRLQLGAPVSRN 451
Qy 281 TSHTVGVWDNAPKTRVCSMTKKEVTEMLRATVNGRYRFAWRELSATFTSNTTFFD 340
Db 452 TPHTVAVNWTPKGRVCTLAKWREIDEMLRDEYQGSYRTVTKTISATFISNTS 511
Qy 341 IILGOCIKRAEAIEOIFRTKYNDSHVKYCHVOYFLALGQFIVAYQPVLSKSLAHM 400
Db 512 IRLGDCATKEAAEADRIYKSKYTHIQGTLETYLARGFLFLAFRPMISNELAKYIN 571
Qy 401 ELMRDNRTEMLDLVNNKHAIIYKKNATSL--SLRLRDIRNAPNRKI-TLDDTTAIK 457
Db 572 ELARSNRTVDLSALLNPSETVQTRGVSQPSNQHRSSRTIEGGIETVNNASLLKTT 631
Qy 458 VOFAMLOFLDHIQTHINDMFSRTATWCELONRELVLWHEGINKINFSATSLGRVA 517
Db 632 VEFAMIOFADYIOAHVNEMLSRITATWCTLONREHVLWTETLKNPGVGVYVSMAL 691
Qy 518 AKMLGDVAAYSSCTAIDAESVTQNSMRVITSTWTCYSRPLVLFPSYGENOQNIQ 577
Db 692 ARLLGDVAVTQCVNISSGHVYIQNSMRVTCSTTCYSRPLVSPRALNDSYIEGG 751
Qy 578 NELLPTLEAVEPCSANHRRYFLGSGYALFENYFVKVDAADIQIASTFVELMLT 637
Db 752 NDLLVERKLIETPCVNNKRYFKGADYVYFEDYAVYRKPVLSEIELISAYVDL 811
Qy 638 REILPLSVYKEELRDVGLDYAEVARNQLHELFYDINKVIEVDN-----YA 687
Db 812 REFPLEVYTRAELEDGLDYSIQRRNLHAKFYDIDISIVRVDNNLVIMRGMA 871
Qy 688 GLOBEGC-----MSTTKDANPNNGOTOLEAARMELTOLINA----- 724
Db 872 GLGDVGAGFGKVLGAASAVISTVSGVSSFLNPPFGALAVGLLILAGIVAFLAY 931
Qy 725 -----KAM---TLASLDYAKIEASLSAYSSEATVNNLNATLEOLKMAK-TN 774
Db 932 LRANPMKALPVITRNKQYAKSPASTAGSDGPDGVDDEBEKLMQAREMIKYMS 991
Qy 775 NQANTDKTTPDNEHNPVLEYA-----KALKT---TLEQRATNLEGSSATYNOI 820
Db 992 EQ-QEHKAMKKNKGPAILTSHLNMALRRRGPKYQRLNNLDSGDDTETNLV 1041

RESULT 13
US-08-413-118-4
; Sequence 4, Application US/08413118
; Patent No. 5688920

```

```

; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-4

Query Match      32.9%: Score 1846; DB 1; Length 1041;
Best Local Similarity 40.4%: Pred. No. 4.4e-117;
Matches 384; Conservative 158; Mismatches 259; Indels 150; Gaps 13;

Qy 18 GTNPSPTQNTVSREVSSVOLSEEE-----STFYLCPPVPVSTVIRLEPPRKCPEPRKAT 73
Db 93 GTPPKPTDPTDMSDREALRASQIEANGPSTFYMCPPSPGVVLEPPRACPDYKLGK 152
Qy 74 EWGEGIALPKENISPKFKVLYKNIQTITWTGTYRQITNRYTDRTPVSEETDL 133
Db 153 NFTGIAVIFKENIAPYKFKANIYKNIIMTVWSSGSYAVTTRNRYTDRVPVKVQETDL 212
Qy 134 IDGRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESR----- 178
Db 213 IDRGMCLSKADYVYRNNYQFTAFDRDEDPRELPLKPSKFNTPQSRGHWYKFKATVYK 272
Qy 179 ----- 178
Db 273 VIVSTANAGSSYTOITNRYADRPVPIVSEITDITDKFGKSSKATYVYRNNHKVFAFNEDK 332
Qy 179 -----AMHTTNETYVWGSPIYRTGTSVNCEIVEEDARSVPYYSYF 220
Db 333 NPQDMLIASKYNSVGSKAHHTTNETYKGAAGFHHSQTSVNCIVEEDARSVPYDSF 392
Qy 221 AMANGDIANISPFYGLSPPEAAEPGYPQDNFQDLSYFSMDLDRKRSKASLPVKRNF 280
Db 393 AISTGDVHNSPFFGLR-DGAHVHTSYSSDRFQOIEGYPIDLDTLRLQLGAPVSRN 451
Qy 281 TSHTVGVWDNAPKTRVCSMTKKEVTEMLRATVNGRYRFAWRELSATFTSNTTFFD 340
Db 452 TPHTVAVNWTPKGRVCTLAKWREIDEMLRDEYQGSYRTVTKTISATFISNTS 511
Qy 341 IILGOCIKRAEAIEOIFRTKYNDSHVKYCHVOYFLALGQFIVAYQPVLSKSLAHM 400
Db 512 IRLGDCATKEAAEADRIYKSKYTHIQGTLETYLARGFLFLAFRPMISNELAKYIN 571
Qy 401 ELMRDNRTEMLDLVNNKHAIIYKKNATSL--SLRLRDIRNAPNRKI-TLDDTTAIK 457
Db 572 ELARSNRTVDLSALLNPSETVQTRGVSQPSNQHRSSRTIEGGIETVNNASLLKTT 631
Qy 458 VOFAMLOFLDHIQTHINDMFSRTATWCELONRELVLWHEGINKINFSATSLGRVA 517
Db 632 VEFAMIOFADYIOAHVNEMLSRITATWCTLONREHVLWTETLKNPGVGVYVSMAL 691
Qy 518 AKMLGDVAAYSSCTAIDAESVTQNSMRVITSTWTCYSRPLVLFPSYGENOQNIQ 577
Db 692 ARLLGDVAVTQCVNISSGHVYIQNSMRVTCSTTCYSRPLVSPRALNDSYIEGG 751
Qy 578 NELLPTLEAVEPCSANHRRYFLGSGYALFENYFVKVDAADIQIASTFVELMLT 637
Db 752 NDLLVERKLIETPCVNNKRYFKGADYVYFEDYAVYRKPVLSEIELISAYVDL 811
Qy 638 REILPLSVYKEELRDVGLDYAEVARNQLHELFYDINKVIEVDN-----YA 687
Db 812 REFPLEVYTRAELEDGLDYSIQRRNLHAKFYDIDISIVRVDNNLVIMRGMA 871
Qy 688 GLOBEGC-----MSTTKDANPNNGOTOLEAARMELTOLINA----- 724
Db 872 GLGDVGAGFGKVLGAASAVISTVSGVSSFLNPPFGALAVGLLILAGIVAFLAY 931
Qy 725 -----KAM---TLASLDYAKIEASLSAYSSEATVNNLNATLEOLKMAK-TN 774
Db 932 LRANPMKALPVITRNKQYAKSPASTAGSDGPDGVDDEBEKLMQAREMIKYMS 991
Qy 775 NQANTDKTTPDNEHNPVLEYA-----KALKT---TLEQRATNLEGSSATYNOI 820
Db 992 EQ-QEHKAMKKNKGPAILTSHLNMALRRRGPKYQRLNNLDSGDDTETNLV 1041

```


Db 932 LRANPKALYPTTRNLKQAKSPASTAGSDPGVDVDFDEKLMQAREMIKYMISLYSAM 991
Qy 775 NQANTDKTTFNEHNPVLEAY---KALKT--TLQRTATNLEGLSSTAYNQI 820
Db 992 EQ-QEHKAMKNGKPAIILTNMLRRRGPKYQRLNNLSDGDDTETNLV 1041

RESULT 15
US-09-232-468A-2
; Sequence 2, Application US/09232468A
; Patent No. 6207165
; GENERAL INFORMATION:
; APPLICANT: AUDONNET et al.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST PORCINE
; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY PATHOLOGIES
; FILE REFERENCE: 454313-2230
; CURRENT APPLICATION NUMBER: US/09/232.468A
; CURRENT FILING DATE: 1999-01-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Pseudorabies virus
US-09-232-468A-2

Query Match 31.7%; Score 1779; DB 3; Length 913;
Best Local Similarity 45.7%; Pred. No. 1.3e-112;
Matches 374; Conservative 130; Mismatches 245; Indels 70; Gaps 19;

Qy 18 GTNSPSTQ--NVTsREVSVQVLSSEESTFYLCPPVGVSTVIRLEPPKCPKPEPRKATEW 75
Db 95 GPSEAPDGEYGDLDARTAYRAA--ATERDRFYVCPFPSPGSGIVVLEPEQACPEYSQGRNF 152
Qy 76 GEGTALLKENISPKYKVTLYKNIIOITTTGTTTQINRYDTNRYDTPVSIETDLID 135
Db 153 TEGIALLFKENIAPKFKAHYIKYKVIIVTTWVGSTYAAITNRTDRVPVQBITVID 212
Qy 136 GKRCSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKNTPESSRAWHHTNETYTVWGSFWI 195
Db 213 RRGKCVSKAEYVRNHNKVTAFDRDENPVEVDLRPSRLNALGTRAWHTNDTYTKIGAAGF 272
Qy 196 YRTGTSVNCIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPEAAAEPMGYPDQNFQ 255
Db 273 YOTGTSVNCIVEEARSVPYDSFALSTGDIVTMSPFYGLR-BSAHGEQIGYAPGRFQ 331
Qy 256 LDSFSDMLDRRKASLPVRKNFLITSHFTVGVWDWAPKTRVCSMTKWKEVTEMLR-ATV 314
Db 332 VEHYPIDLDSRLRASESVTRNFTPHFTVAMDWAPKTRRVCSLAKWREAEEMTRDTR 391
Qy 315 NGRYREWARELSATFISNTTFDPNRIILGOCIKREAEAELEQIFRTKYNDSHVKG-HV 373
Db 392 DGSFRTSRALGASFVSDVTQDLQRVHLGDCVLEKASEALDAIRRYNSTHVHLAGDRP 451
Qy 374 QYFLALGGFVAYQVLSKSLAHMYLRELNRDNTDEMLDLVNNKHAIYKKNATSLSLRL 433
Db 452 EVYLARGGVVAFRPLISNELAQYAREL-----ERLGLAG---VVGPAAPAAARR 501
Qy 434 RD-----TRNAPNRKITDDTTAKSTSSVQFAMQLQFLYDHIQTHINDMFRIA 482
Db 502 RSPGPACTPEPPAYNGTGHRLIT-----TGSAEFARLQFTYDHIQAHVNDMLGRIA 552
Qy 483 TAWCELQNRRELKINWHEGKINPSATASATLGRVRAAKMGDVAASVSCATAIDAESVTQ 542
Db 553 AAWCELQNKORTLHSEMSRLNPSAVATAALQQRVCAKMLGDVMAISRCVEVRG-GVYVQN 611
Qy 543 SMRVITSTNTCYSRPLVLFYSYGENQIQQLGENNELPTLEAVEPCSANHRRYFLFGS 602
Db 612 SMRVPGERGTCYSRPLVTFEH-NGTVIEGQGLGDDNELLISRODLIEPCTGNHRRYFKLGS 670
Qy 603 GYALFENYFVKWDAADIQASTFVELNLTLLDREILPLSVYTKKEELRDVGVDYAEV 662
Db 671 GYVYEDYNVVRWVEVP--ETISTRVNTLTLLEDREPLFLEVYTREELADTGLLDYSEI 728

Qy 663 ARNQLHELKFXDINKVIEVDTN-----YAGLOEFGC-----MSTTKKDAMPNN 706
Db 729 QRRNQLHALKFXDIDRVVVKVDHNVVLLRGIANFFQGLGDVGAAGVKVVLGATGAVISAVG 788
Qy 707 GQTOLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEATVNNNLNATLEQLKMA 766
Db 789 GWSF-----LSNPFGALAIGLLVL---AGLVAAF-LAYRHISRLRRNPMKALYPV-TT 837
Qy 767 KTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEOR 805
Db 838 KTLKEDGVDEGDVDEAKLDQARDMI--RYMSIVSALEQQ 874

Search completed: October 8, 2003, 17:09:38
Job time : 34.5244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:10 ; Search time 88.2345 Seconds
(without alignments)
1953.624 Million cell updates/sec

Title: US-09-147-052-4
Perfect score: 5619
Sequence: 1 MHYFRNCIFLIYLIGTN.....SSNENNADKIPGYRRPGTFL 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5608	99.8	1086	18	AAW36051
2	3601	64.1	865	14	AAW30169
3	2015	35.9	456	18	AAW36050
4	1968	35.0	879	16	AAW77024
5	1913	34.0	615	15	AAW63230
6	1859	33.1	610	15	AAW63229
7	1779	31.7	868	13	AAW22615
8	1779	31.7	913	19	AAW68404
9	1765.5	31.4	845	6	AAW50035
					N-terminal sequence

10	1746	31.1	933	13	AAW27807	Bovine herpes viru
11	1746	31.1	933	14	AAW41343	Bovine herpesvirus
12	1743	31.0	943	21	AAW32470	DNA encoding felin
13	1743	31.0	943	22	AAW50113	Feline herpesvirus
14	1739.5	31.0	891	17	AAW92746	B virus gB glycopr
15	1739.5	31.0	891	19	AAW70293	Simian herpesvirus
16	1736	30.9	904	23	AAE17812	Herpes simplex vir
17	1735	30.9	854	19	AAW72113	Herpes simplex vir
18	1735	30.9	904	14	AAW41778	HSV-2 strain SB5 C
19	1735	30.9	904	17	AAW00376	Glycoprotein B (gB
20	1735	30.9	904	18	AAW34552	HSV-2 glycoprotein
21	1735	30.9	904	19	AAW72193	Herpes simplex vir
22	1735	30.9	904	22	AAW74442	HSV-2 strain SB5 C
23	1732	30.8	854	8	AAW70347	Herpes simplex vir
24	1726	30.7	904	9	AAW80914	Varicella-zoster v
25	1720	30.6	907	8	AAW71136	Sequence of Herpes
26	1719.5	30.6	885	17	AAW92747	SAB virus gB glyco
27	1705.5	30.4	904	17	AAW00375	HSV-1 glycoprotein
28	1705.5	30.4	904	22	AAW74441	Herpes simplex vir
29	1704.5	30.3	795	19	AAW72062	HSV-2 strain SB5 C
30	1703.5	30.3	903	8	AAW70426	Recombinant herpes
31	1702.5	30.3	904	12	AAW14665	HSVgB polypeptide.
32	1700.5	30.3	903	7	AAW60244	Herpes simplex vir
33	1699.5	30.2	903	8	AAW71135	Herpes simplex vir
34	1696.5	30.2	904	14	AAW41779	Glycoprotein B (gB
35	1696.5	30.2	905	9	AAW80915	Sequence of Herpes
36	1694.5	30.2	973	8	AAW70769	Glycoprotein B of
37	1690.5	30.1	904	18	AAW34553	Herpes simplex vir
38	1690.5	30.1	973	12	AAW14680	HSV surface antige
39	1682	29.9	694	12	AAW14666	Truncated HSVgB po
40	1615	28.7	368	14	AAW44493	Mycoplasma gallise
41	1615	28.7	368	15	AAW63227	Mycoplasma gallise
42	1607	28.6	368	16	AAW76955	Mycoplasma gallise
43	1601	28.5	928	16	AAW77399	BHV1 gI glycoprote
44	1576	28.0	903	6	AAW50312	Herpes simplex vir
45	1344.5	23.9	932	19	AAW44947	Bovine herpesvirus

ALIGNMENTS

RESULT 1

AAW36051
ID AAW36051 standard; Protein; 1086 AA.

AC AAW36051;

XX

XX 15-JUL-1998 (first entry)

XX Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.

DE Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

KW antigen; vaccine; poultry.

XX

OS Chimeric - Marek's disease gammaherpesvirus.

OS Chimeric - Mycoplasma gallisepticum.

XX

XX Key Location/Qualifiers

FT Region 1..672

FT Region /note- "derived from Marek's disease virus gB protein"

FT Region 693..1086

FT Region /note- "derived from M. gallisepticum antigen"

XX W09736924-A1.

XX

XX 09-OCT-1997.

XX

XX 28-MAR-1997; 97WO-JP01084.

XX

XX 29-MAR-1996; 96JP-0103548.

XX (JAPG) NIPPON ZEON KK.

XX

PI Saito S, Tsuzaki Y, Yanagida N;
 DR WPI; 1997-503045/46.
 DR N-PSDB; AAT96596.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;

Query Match 99.8%; Score 5608; DB 18; Length 1086;
 Best Local Similarity 99.78; Pred. No. 0;
 Matches 1083; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MHVFRNCIFFLIVILYGTNSPSTQNTSREVSVVQLSEESTFYLCPPPVGTVIRL 60
 Db 1 MHVFRNCIFFLIVILYGTNSPSTQNTSREVSVVQLSEESTFYLCPPPVGTVIRL 60
 Qy 61 EPPKCEPRKATWEGGIAILLKENISPKFKVLYKNIQTWTGTYRITNRYT 120
 Db 61 EPPKCEPRKATWEGGIAILLKENISPKFKVLYKNIQTWTGTYRITNRYT 120
 Qy 121 DRTPVSEETDLDIGRGRSSKARYLRNNVYVAFDRDAGEKQVLLKPSKENTPESRAW 180
 Db 121 DRTPVSEETDLDIGRGRSSKARYLRNNVYVAFDRDAGEKQVLLKPSKENTPESRAW 180
 Qy 181 HTTNETYVWGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSPPFGLSPPE 240
 Db 181 HTTNETYVWGSPWYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANSPPFGLSPPE 240
 Qy 241 AAAPMGYPQDNFKQDLSYFMDLDRKRKASLPVKNFLITSHPTGWDWAPKTRVCSM 300
 Db 241 AAAPMGYPQDNFKQDLSYFMDLDRKRKASLPVKNFLITSHPTGWDWAPKTRVCSM 300
 Qy 301 TKWKEVTEMLRATVNGRYRFAWELSATFISNTTEPDNRIILGQCIRKRAEAIEQIFR 360
 Db 301 TKWKEVTEMLRATVNGRYRFAWELSATFISNTTEPDNRIILGQCIRKRAEAIEQIFR 360
 Qy 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNTEMLDLVNNKHA 420
 Db 361 TKYNDSHVKVGHVQYFALGGFIVAYQPVLSKSLAHMYLRELMDNRNTEMLDLVNNKHA 420
 Qy 421 IYKKNATSLRLRDRIRNAPNRKITLDDTTAISTSSVQFAMLOFLYDHIQTHINDMFSR 480
 Db 421 IYKKNATSLRLRDRIRNAPNRKITLDDTTAISTSSVQFAMLOFLYDHIQTHINDMFSR 480
 Qy 481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAVSSCTAIDAESVTL 540
 Db 481 IATAWCELQNLRELVLWHEGKINPSATASATLGRRAAKMLGDVAVSSCTAIDAESVTL 540
 Qy 541 QNSMRVITSTNTCYSRPLVLFSGENOGNTOGQGENNELPTLEAVEPCSAHRRYFLF 600
 Db 541 QNSMRVITSTNTCYSRPLVLFSGENOGNTOGQGENNELPTLEAVEPCSAHRRYFLF 600
 Qy 601 GSGYALFENYFVKMVDAAADIIQASTFVELNLTLLEDREILPLSVYTKBELROGVLDYA 660
 Db 601 GSGYALFENYFVKMVDAAADIIQASTFVELNLTLLEDREILPLSVYTKBELROGVLDYA 660
 Qy 661 EVARRNQLHELKFDYDINKVIEDVTNYAGLQEFCCMSITTKDANPNNGOTOLEAARMELTD 720
 Db 661 EVARRNQLHELKFDYDINKVIEDVTNYAGLQEFCCMSITTKDANPNNGOTOLEAARMELTD 720
 Qy 721 LINAKAWTLASLODYAKIEASLSAYSEATVNNNLNATLEQLKMAKTINLESAINOANTD 780

Db 721 LINAKAWTLASLODYAKIEASLSAYSEATVNNNLNATLEQLKMAKTINLESAINOANTD 780
 Qy 781 KTTFDNEHPNLEVEAYKALKTTLEQATNLEGLSSTAYNOIRNNLDLNKASSLITKTLD 840
 Db 781 KTTFDNEHPNLEVEAYKALKTTLEQATNLEGLSSTAYNOIRNNLDLNKASSLITKTLD 840
 Qy 841 PLNGGTLSDSNEITANKNINNTLSTINEQKTNADALSNSFTKKYVIONNEQSFVGTFTNA 900
 Db 841 PLNGGTLSDSNEITANKNINNTLSTINEQKTNADALSNSFTKKYVIONNEQSFVGTFTNA 900
 Qy 901 NVQPSNTSFVAFSADVTPVNYKYARTVWNGDESSRLANTNSITDYSWISLAGTNTK 960
 Db 901 NVQPSNTSFVAFSADVTPVNYKYARTVWNGDESSRLANTNSITDYSWISLAGTNTK 960
 Qy 961 YQFSFNTSGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVVEFATSTANNTANTPT 1020
 Db 961 YQFSFNTSGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVVEFATSTANNTANTPT 1020
 Qy 1021 PAVDEIKVAKIVLSGLRFGQNTIELSVPTGSGNNKVPAMIGNIYLSNENNADKIPGYR 1080
 Db 1021 PAVDEIKVAKIVLSGLRFGQNTIELSVPTGSGNNKVPAMIGNIYLSNENNADKIPGYR 1080
 Qy 1081 RPTGFL 1086
 Db 1081 RPTGFL 1086

RESULT 2
 AAR30169
 ID AAR30169 standard; Protein; 865 AA.
 XX AAR30169;
 AC AAR30169;
 XX
 DT 25-MAR-2003 (updated)
 DT 07-MAY-1993 (first entry)
 XX
 XX Marek's Disease Virus glycoprotein B homologue of HSV.
 XX Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
 KW Herpes Simplex Virus.
 XX Marek's Disease Virus.
 OS
 XX EP520753-A1.
 XX 30-DEC-1992.
 XX
 XX 24-JUN-1992; 92EP-0305775.
 XX
 XX 28-JUN-1991; 91US-0722860.
 PR 30-DEC-1991; 91US-0803633.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (USDA) US SEC OF AGRIC.
 XX
 XX Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
 WPI; 1993-001546/01.
 XX
 XX Recombinant fowl pox virus contg. Marek's disease virus antigen
 PT gene - used to produce cell-free vaccine against Marek's disease
 PT virus
 XX
 XX Example 2; Page 15-19; 30pp; English.
 XX
 CC The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
 CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-Sali
 CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
 CC from K3 fragment were ligated with EcoRI, Sali digested pUC18. The
 CC sequence of the putative MDV gBh was determined by sequencing a set
 CC of deletion mutants. The nucleotide and amino acid sequences were
 CC found to be identical with the published sequences of the gBh of

CC RBIB strain of MDV (Ross et al., J. Gen. Virol., 70:1789-1894, 1988).
 CC A fragment contg. the entire coding region of MDV g8h was inserted
 CC into pNZ1729R (see AAQ34774-Q34778) to produce a recombinant FPV/MDVgBh
 CC virus for immunising chickens. See also AAQ34780-Q34781.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 865 AA;

Query Match 64.1%; Score 3601; DB 14; Length 865;
 Best Local Similarity 85.9%; Pred. No. 5.9e-219;
 Matches 719; Conservative 16; Mismatches 60; Indels 42; Gaps 8;

QY 1 MHYFRNCIFELVILYGNSSPSTONVTSREVSSVQLSEESTFYLCPPPGVSVIRL 60
 DB 1 MHYFRNCIFELVILYGNSSPSTONVTSREVSSVQLSEESTFYLCPPPGVSVIRL 60
 QY 61 EPPKCPKPEKATWEGEGIAILFKENISPYKFKVTLYYKNIQTTTWTGTYQINRYT 120
 DB 61 EPPKCPKPEKATWEGEGIAILFKENISPYKFKVTLYYKNIQTTTWTGTYQINRYT 120
 QY 121 DRTPVSIETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 DB 121 DRTPVSIETDLDIGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
 QY 181 HTTNETYTVWGSPIYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANISPYGLSPPE 240
 DB 181 HTTNETYTVWGSPIYRTGTSVNCIVEEMDARSVPYSYFAMANGDIANISPYGLSPPE 240
 QY 241 AAAPMGYPQDNFKQLDSVFSMDLKRKASLPVKRNFLTITSHFTVGDWAPKTRVCSM 300
 DB 241 AAAPMGYPQDNFKQLDSVFSMDLKRKASLPVKRNFLTITSHFTVGDWAPKTRVCSM 300
 QY 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTFDPNRIILGQCCKREAEAAIQIFR 360
 DB 301 TKWKEVTEMLRATVNGRYRPMARELSATFISNTTFDPNRIILGQCCKREAEAAIQIFR 360
 QY 361 TKYNDSHVKVGHVQYFALGGFIVAYOPVLSKSLAHMYLRELRDNRDTEMLDYNKHA 420
 DB 361 TKYNDSHVKVGHVQYFALGGFIVAYOPVLSKSLAHMYLRELRDNRDTEMLDYNKHA 420
 QY 421 IYKKNATSLRLRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 DB 421 IYKKNATSLRLRDIRNAPNRKITLDDTTAISTSSVQFAMQLFYDHIQTHINDMFSR 480
 QY 481 IATAWCELQNLRELVLVHHEGIIKINPSATASATLGRVAAKMLGDVAAYSSCTAIDAEVTL 540
 DB 481 IATAWCELQNLRELVLVHHEGIIKINPSATASATLGRVAAKMLGDVAAYSSCTAIDAEVTL 540
 QY 541 QNSMRVITSTNTCYSRPLVLSYGENQNIQOGQGENNELLPLEAVEPCSANRRYFLF 600
 DB 541 QNSMRVITSTNTCYSRPLVLSYGENQNIQOGQGENNELLPLEAVEPCSANRRYFLF 600
 QY 601 GSGYALFENYFKVMYDAADIIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660
 DB 601 GSGYALFENYFKVMYDAADIIQIASTFVELNLTLLEDREILPLSVYTKKEELRDVGVLDA 660
 QY 661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKADPNNGQTL-BAARM 716
 DB 661 EVARRNQLHELKPYDINKVIEVDNTYA---GLQEFQCMSTTKKADPNNGQTL-BAARM 716
 QY 717 ELTDLLINAKMTLASLQDYAK-----TEASLSAYSEAEVTN---NNLNA--- 758
 DB 709 VVGAAGAIVSTISSGSAFMSNPFAGALITGLIIAGLVAAFLAYRVNKLKSNPMKALYP 768
 QY 759 -TLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEA-----YKALKTTLEQRAATNL 809
 DB 769 MTEVLKAQATRELHGEESDDELTSTIDER--KLEAREMIKYMALVSAERHEKKL 823

RESULT 3
 AAW36050
 ID AAW36050 standard; Protein; 456 AA.
 XX

AC AAW36050;
 XX
 DT 15-JUL-1998 (first entry)
 XX
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.
 XX
 KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 XX antigen; vaccine; poultry.
 KW Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 1..64
 FT /note= "derived from Marek's disease virus gB protein"
 FT Region 65..456
 FT /note= "derived from M. gallisepticum antigenic protein"
 XX
 PN W09736924-A1.
 XX
 PD 09-OCT-1997.
 XX
 XX 28-MAR-1997; 97WO-JP01084.
 XX
 PR 29-MAR-1996; 96JP-0103548.
 XX
 XX (JAPG) NIPPON ZEON KK.
 XX
 PI Saito S, Tsuzaki Y, Yanagida N;
 XX
 XX WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; Sipp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 CC
 SQ Sequence 456 AA;

Query Match 35.9%; Score 2015; DB 18; Length 456;
 Best Local Similarity 100.0%; Pred. No. 4.3e-119;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 691 EFGCMSITKKNPNQTOLEAARMBELTDLINAKAMTSLASLODYAKIEASLSAYSEAE 750
 DB 61 EFGCMSITKKNPNQTOLEAARMBELTDLINAKAMTSLASLODYAKIEASLSAYSEAE 120
 QY 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAATNLE 810
 DB 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAATNLE 180
 QY 811 GLSTAYNQIRNNLVLYNKLKASLITKTLPLNGGILLDSNETTANKNNINWLTSTNEQ 870
 DB 181 GLSTAYNQIRNNLVLYNKLKASLITKTLPLNGGILLDSNETTANKNNINWLTSTNEQ 240
 QY 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKVARRTVWN 930
 DB 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKVARRTVWN 300
 QY 931 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 990
 DB 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFSGNYGPGSTGYLYFPYKLVKAADANNV 360
 QY 991 GLQYKLNNGNVQVFEATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050

```

|||||
361 GLOKLNNGNQVEFATSTANNNTANPTFAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
QY 1051 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 1086
|||||
421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456

RESULT 4
AAR77024
ID AAR77024 standard; Protein; 879 AA.
AC AAR77024;
DT 20-APR-1996 (first entry)
DE Canine herpesvirus gB homologue.
KW CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;
KW canarypox virus; ALVAC; VCP320; dog; puppy; immunisation; antigen;
KW vaccine.
OS Canine herpesvirus.
XX
XX Key Location/Qualifiers
XX FT Region 725..741
XX FT /label= Transmembrane_region
XX FT 747..771
XX FT /label= Transmembrane_region
XX
XX WO9526751-A1.
XX
XX 12-OCT-1995.
XX
XX 30-MAR-1995; 95WO-US03982.
XX
XX 29-MAR-1995; 95US-0413118.
XX 30-MAR-1994; 94US-0220151.
XX (VIRO-) VIROGENETICS CORP.
XX
XX Limbach KJ, Paoletti E;
XX WPI; 1995-366131/47.
XX N-PSDB; AAT01402.
XX
XX Nucleic acids encoding canine herpes virus (CHV) gB, gC and gD
XX glyco:proteins - also glyco:proteins and vectors, for the
XX immunisation of neonatal puppies and adult dogs against CHV
XX
XX Example 3; Fig 1A-G; 241pp; English.
XX
XX A canine herpesvirus (CHV) protein has a predicted amino acid
XX sequence (AAR77024) that shows significant homology with the gB
XX glycoprotein of numerous herpesviruses. It is the product of a
XX gene (see AAT01402) isolated from CHV genomic DNA using a probe contg.
XX the feline herpesvirus gB, gC and gD genes. CGV glycoprotein gC and
XX gD homologues (AAR77025-26) have also been obt'd. These glycoproteins,
XX including recombinant glycoproteins expressed from attenuated
XX recombinant virus vectors, e.g. ALVAC recombinant VCP320 (see
XX AAT01406), can be used in antigenic, immunological or vaccine
XX compositions to protect puppies and adult dogs against CHV.
XX
XX Sequence 879 AA;
XX
XX Query Match 35.0%; Score 1968; DB 16; Length 879;
XX Best Local Similarity 45.8%; Pred. NO. 1.1e-115;
XX Matches 387; Conservative 153; Mismatches 253; Indels 52; Gaps 9;
XX
XX 21 SSPSTQNTSREVSSVOL-SEESTFYLCPPPVGVSTVIRLEPPRKCPKPRKATWEGGI 79
XX 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1:::
XX 36 STPKPTSDIREILRESQIESDDTSTFYMCPPPSGSLVRLPPRACPNYKLGKNFTBGI 95
XX 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1::: 1:::

```

```

QY 80 AILFKENISPKFKVTLYYKNIQTTWTGTTTQITNRYTDRTPVSEIEETDLDISKGR 139
DB 96 AVIFKENISPKFKANIYYKNIITTVMSGSTYAITNRYTDRVPVIGVPEITELDRRGM 155
QY 140 CSSKARYLRNVVYEAFFDRDAGEKOVLLKPKSKENTPESRAWHHTNETTYVNGSPWIVRTG 199
DB 156 CLSKADYIRNNYEFTAFDKDDEPREVHLKPSKFNTPGSRGWHVTNDYTTKGGSGFYHSG 215
QY 200 TSVNCIIVEMDARSVPFYSFAMANGDIANTISPFYGLSPPEAAAEPMGYPODNFKQLDSY 259
DB 216 TSVNCIIVEVDARSVPYDSEFAISTGDIHMSPEFGLR-DGAHTEYISYSTDRQQIGSY 274
QY 260 FSDMDKRRKASLPVKRNFLLTSHFTVGVNDWAPKTRVCSMTKWEVTEMLRATVNGRYR 319
DB 275 YPIDLTRQLQAGPVSRRNFLTTOHTVAVNNVPVIREVCTTAKVREIDEIRDEYKGSYR 334
QY 320 FMARELSATFISNTTEFPNRIILGOCIKREAAEAOIOPRTKYNDSHVKGVOYFLAL 379
DB 335 FTKASISATFISDTYQFIDRVKLSDCAKRAIRAIADKIYKKKYNKHTIQTGELETYLR 394
QY 380 GGFIVAYQPVLSKSLAHMYLRELMRDNRTDEMLDVNNKHAIYKKKNTLSRLRRDIRNA 439
DB 395 GGFIIAFRPMISNELAKLYINELVRSNRTVDLSILLNPVSRGARKRRSVEENKSRNI 454
QY 440 PNKTIITDDTTAIKSTSSQFAMQLQFLYDHTIOTINDMFSTRATACELQNRVLVHREG 499
DB 455 EGGIENVNNTSTIIKTSSVHFAMLOFAYDHIQSHVNEMLSRISATACNLQNKERTLNEV 514
QY 500 IKINPSATASATLGRVAAKMLGDVAVSSCTAIDAESVTQLQNSMRVTTSTNTCYSRPLV 559
DB 515 MKNLPTSVASVAMDORVSARMGLDVLAVTQCVNISGSSVFIQNSMRVLGTTTCYSRPLI 574
QY 560 LFSYGENQGN-IOQLGENNELLTLEAVEPCSNHRRYFLFGSGYALFENYFVKWVDA 618
DB 575 SFKALENSTNYIEQGLGENNELVERKLIEPCTANHKRYEFGADYVYFENYAYVRKVP 634
QY 619 ADIOIASFVELNLTLLEDRILPLSVYTKKEELRDGVLDVAEVARRNQLHKLAFYDINK 678
DB 635 NEIEMISAYVDNLTLLEDRFLPLEVYTRAELEDTGLLDYSEIQRRLHALFKFYDIDS 694
QY 679 VIEYDTNYAGLOEFGCMSITKKDANPNNGOTQLEAARMELFDLINAKAMTILASLDYAK- 737
DB 695 VVKVDNNVVIMRGIANFQGLGDVGAGFGKVLGAA-----NAVITATVSGVSSFLNN 746
QY 738 -----TEASLSSAYSEATVNNNLNATLEOLKMAKT-NLESAINQANTDKTTFDN 786
DB 747 PFGALAVGLLILLAGLFAFLAYRYVSKLSNPKALYPVTTKNLKESYKNGNSGNSDGE 806
QY 787 EHPNLVE-----AYKALATTLEOR-----ATNLEGLS-----STAY 817
DB 807 ENDDNIDEKLOQAKEMIKYNSLYSAMEQOEKKAIRKNSGPALLASHITNLSLKHGPKY 866
QY 818 NQIRN 822
DB 867 KRLKN 871
XX
XX RESULT 5
XX AAR63230
XX ID AAR63230 standard; Protein; 615 AA.
XX AC AAR63230;
XX XX
XX 25-MAR-2003 (updated)
XX DT 23-JUN-1995 (first entry)
XX DE
XX DE Mycoplasma gallisepticum antigen (UM-67).
XX XX Mycoplasma avipox virus; live vaccine; mycoplasma antigen.
XX XX Mycoplasma gallisepticum.
XX XX
XX FH Key Location/Qualifiers

```



```

FT Protein      1..615
TT /note= "Trp residues correspond to TGA codons"
XX WO9423019-A1.
XX 13-OCT-1994.
XX 31-MAR-1994; 94WO-JP00541.
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX WPI: 1994-333181/41.
XX N-PSDB; AAQ77857.
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX Claim 4; Page 87-91; 123pp; Japanese.
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX ORF encodes an antigenic polypeptide (AAQ77857). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 615 AA;
    Query Match      34.0%; Score 1913; DB 15; Length 615;
    Best Local Similarity 95.9%; Pred. No. 1.9e-112;
    Matches 377; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
    QY 684 TNYAGLQFEGCMSITTKDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLS 743
    DB 17 TSFSLSIGISSCSITTKDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLS 76
    QY 744 SAYSAETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803
    DB 77 SAYSAETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 136
    QY 804 QRATNLEGLSTAYNQIRNLDYLNKASSLITKTLDPNGGTLDSNEITANKNINNT 863
    DB 137 QRATNLEGLSTAYNQIRNLDYLNKASSLITKTLDPNGGTLDSNEITANKNINNT 196
    QY 864 LSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKY 923
    DB 197 LSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKY 256
    QY 924 ARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYCPSTGYLYFPYKLVK 983
    DB 257 ARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYCPSTGYLYFPYKLVK 316
    QY 984 ADANNVGLQYKLNGNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTI 1043
    DB 317 ADANNVGLQYKLNGNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTI 376
    QY 1044 ELSVPTGEGNMKVPAMIGNIYLSNENNADKI 1076
    DB 377 ELSVPTGEGNMKVPAMIGNIYLSNENNADKI 409
    RESULT 6
    ID AAR63229
    XX AAR63229 standard; Protein; 610 AA.

```

```

AC AAR63229;
XX 25-MAR-2003 (updated)
XX 23-JUN-1995 (first entry)
XX Mycoplasma gallisepticum antigen (UM-66).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
XX Key Location/Qualifiers
XX Protein 1..610
XX /note= "Trp residues correspond to TGA codons"
XX WO9423019-A1.
XX 13-OCT-1994.
XX 31-MAR-1994; 94WO-JP00541.
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX WPI: 1994-333181/41.
XX N-PSDB; AAQ77856.
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX Claim 4; Page 78-81; 123pp; Japanese.
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX ORF encodes an antigenic polypeptide (AAQ77856). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 610 AA;
    Query Match      33.1%; Score 1859; DB 15; Length 610;
    Best Local Similarity 93.4%; Pred. No. 4.8e-109;
    Matches 366; Conservative 13; Mismatches 13; Indels 0; Gaps 0;
    QY 684 TNYAGLQFEGCMSITTKDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLS 743
    DB 17 TSFSLSIGISSCSITTKDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLS 76
    QY 744 SAYSAETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 803
    DB 77 SAYSAETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLE 136
    QY 804 QRATNLEGLSTAYNQIRNLDYLNKASSLITKTLDPNGGTLDSNEITANKNINNT 863
    DB 137 QRATNLEGLSTAYNQIRNLDYLNKASSLITKTLDPNGGTLDSNEITANKNINNT 196
    QY 864 LSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKY 923
    DB 197 LSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKY 256
    QY 924 ARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYCPSTGYLYFPYKLVK 983
    DB 257 ARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYCPSTGYLYFPYKLVK 316
    QY 984 ADANNVGLQYKLNGNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTI 1043

```


KW	suid herpesvirus 1; swine influenza virus; hog cholera virus; vector.
KW	porcine respiratory and reproductive syndrome virus; glycoprotein; SIRSv.
KW	swine infertility and respiratory syndrome virus;
KW	Actinobacillus pleuropneumoniae.
XX	
OS	Pseudorabies virus.
XX	
PN	FR2751224-A1.
XX	
PD	23-JAN-1998.
XX	
PF	19-JUL-1996; 96FR-0009338.
XX	
PR	19-JUL-1996; 96FR-0009338.
XX	
PA	(INMR) RHONE MERIEUX SA.
XX	
DR	WPI; 1998-112824/11.
XX	N-PSDB; AAV49291.
DR	
XX	Multi-valent polynucleotide vaccines against porcine pathogens -
PT	consist of at least 3 plasmids able to express protective antigens
PT	from specified viruses
XX	
PS	Example 8; Fig 2; 63pp; French.
XX	
CC	The invention relates to a multivalent vaccine for protecting pigs
CC	against several pathogens, especially pathogens associated with
CC	respiratory and digestive diseases. The pathogens are especially
CC	selected from Aujeszky's disease virus, swine influenza virus (SIV),
CC	porcine respiratory and reproductive syndrome virus (PRRSV), hog
CC	cholera virus (HCV) and Actinobacillus pleuropneumoniae. The vaccines
CC	are preferably composed of polynucleotide sequences encoding 3 antigens,
CC	all as part of vectors. This sequence represents the Aujeszky's disease
CC	virus strain NIA3 glycoprotein gB. The coding sequence was subcloned
CC	into the plasmid pVR1012 to generate plasmid pAB090 for use in the
XX	vaccine.
XX	
SQ	Sequence 913 AA;

Db	502	RSPGPAGTPEPPAVNGTGH	LRIT-----TGSAEFARLQFTYDHIQAHVNDMLGRIA	552	
Qy	483	TAWCELONRELVLHHEG	IKNPSATATLGRRRYAAKMLGDVAAVSSCTADAESVTQLON	542	
Db	553	AANCELQNKORTLWSE	RLNPSAVATAALGQRVCARMLGDVAISRVCVEVRG-GVYVON	611	
Qy	543	SMRVITNTCYSRPLF	SYSGENQGIQOLGENNELLP	LEAVEPCSANHRRYFLFGS	602
Db	612	SMRVPGERGTCYSR	PLTYFEH-NCTGVTEGQ	LDNELLISRDLEPCTGNHRRYFKLGS	670
Qy	603	GYALFENYFVKWDA	ADIQIASTFVELNLTLLEDREILP	SVYTKKELRDVGVLDAEV	662
Db	671	GYVYEDYNTVRW	VEVP--ETISTRVTINLTLEDREFL	PEVTVTREETLADTGLDYSEI	728
Qy	663	ARRNQLHELKIFYD	INKVLEVDTN-----YAGLQEF	GC-----MSITKKDANPNN	706
Db	729	QRRNQLHALKFYD	IDRVKVDHNVLLRGIANFFQGLGDVGA	AAVKVVGATGAVISAVG	788
Qy	707	GOTOLEAARMELTD	LINAKAMTSLASLDYAKIEASLSAYSE	AEETVNNNNLNATLEQLKMA	766
Db	789	GWYSF-----	LSNPFGALAGLLVL---AGLVAA	F-LAYRHSILRLRNPKALYPV-TT	837
Qy	767	KTNLIESAINANT	OKTTFDNEHPNLV	RAYKALKTTLEOR	805
Db	838	KTLEDGVDEGDV	DEAKLDQARDMI--RYMSIV	SALEQQ	874
RESULT 9					
AAP50035					
ID	AAP50035 standard; Protein; 845 AA.				
AC	AAP50035;				
XX					
DT	25-MAR-2003 (updated)				
DT	07-SEP-1991 (first entry)				
XX					
DE	N-terminal sequence of the pseudorabies virus (PRV) gII protein.				
KW	Sub-unit vaccine; antigen; immunoreactive determinant.				
XX					
OS	Pseudorabies virus (PRV).				
XX					
FH	Location/Qualifiers				
FT	Modified-site	151..153			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	261..263			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	441..443			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	516..518			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	573..575			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	633..635			
FT	FT	/label=	potential glycosylation site		
FT	Modified-site	697..699			
FT	FT	/label=	potential glycosylation site		
XX					
PN	EPI62738-A.				
XX					
PD	27-NOV-1985.				
XX					
PF	09-APR-1985; 85EP-0400704.				
XX					
PR	03-APR-1985; 85US-0719773.				
PR	09-APR-1984; 84US-0598073.				
XX					
PA	(MOLE-) MOLECULAR GENETICS INC.				
XX					
PI	Robbins AK, Watson RJ, Enquist LW;				
XX	WPI; 1985-298083/48.				
DR					

(MOLE-) MOLECULAR GENETICS INC.

Robbins AK, Watson RJ, Enquist LW;

XX
DR WPI; 1985-298083/48.

DR N-PSDB; AAN50036.
 XX Prodn. of Pseudorabies virus sub-unit vaccines - useful for
 PT conferring protection against the virus infections and for
 PT diagnosis esp. with sheep sera
 XX
 PS Example; Fig 13; 153pp; English.
 XX
 CC The inventors claim a recombinant vector comprising a DNA sequence
 CC coding for an immunoreactive and antigenic determinant of a
 CC Pseudorabies virus protein, and a polypeptide having an
 CC immunoreactive and antigenic determinant of a Pseudorabies virus
 CC protein. Vaccines contg. the polypeptide are obtd. economically and
 CC in large amts., for use in conferring protection against
 CC Pseudorabies virus.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 845 AA;

Query Match 31.4%; Score 1765.5; DB 6; Length 845;
 Best Local Similarity 50.1%; Pred. No. 6.3e-103;
 Matches 351; Conservative 111; Mismatches 188; Indels 51; Gaps 13;

QY 18 GTNSSPSTQ--NVTSEVSSVQLSEESTFYLCPPPGVSTVIRLEPPKCPPEPKATEW 75
 DB 95 GSEAPDGEYGLDARTAA--ATERDRFVCPPPSGSTVIRLEPEQACPEYSQGRNF 152
 QY 76 GEGIAIFKENISPKFKVTLTKYKNIQTWTGTYRQITNRYTDRTPVSEIEITDLID 135
 DB 153 TEGIAVLFKENIAHFKAHYIKVIVTVMSGSTYAITNRTFDRVPVQETIDVID 212
 QY 136 GGRGSSKARYLRNNVYFAFDORAGEKQVLLKPSKFNTPESRAWHHTTNETYVMSPI 195
 DB 213 RGKCVSKAEYVRNNHKVTAFDRENPEVDLRPSRLNALGTRGWHHTNDYTKIGAAGF 272
 QY 196 YRTGTSVNCIVEEMDARSVEPYSPAMANGDINTANSPFYGLSPPEAAEPYQDNFKQ 255
 DB 273 YHTGTSVNCIVEEARSVPYDPSFALSGDVIYMSPPFGLR-EGAHGHIYAGPFGQ 331
 QY 256 LDSYFMDLDRKRLASLPVKRNFLTSHFTVGMWAPKTRVCSMTKWKEVTEMLR-AIV 314
 DB 332 VEHYTPIDLDLSRLASRSEVTRNFLTPTFTVAMDWAPKTRRVCSLAKWREAEEMTRDET 391
 QY 315 NGRYRFMARELSATISNTEPDNRITLIGQIKREAEAAIQIFRTKYNDSHVKVQ-HV 373
 DB 392 DGSFRTSALGASFVSDVTQDLQVHLGDCVLRSEASEAIDAIYRRYRNSHVLGDRP 451
 QY 374 QYFLALGGFIVAYQVPLSKSLAHMYLRELMDRNRDEMLDLVNNKHAIYKKNATSLSLR 433
 DB 452 EYVLARGGVFAFRPLISNELAQLYAREL-----ERGLAG---VVGPAAPAAARRAR 501
 QY 434 RD-----IRNAPNRKITDTPAITSSTSSQVAMQLQDHYDHIQTHINDMFSRIA 482
 DB 502 RSPGPAGTPEPPAVNGTGLRIT-----TGSAEFAQLQFTYDHIQAHVNDMLGRIA 552
 QY 483 TAWCELOHRELVLWHEGKINPASFATSLGRVAAMGLGDAVVAVSSCTAIDAESVTQON 542
 DB 553 AAWCELOHRELVLWHEGKINPASFATSLGRVAAMGLGDAVVAVSSCTAIDAESVTQON 611
 QY 543 SMRVITSTNTCSRLPLFSYGENOGNTOGOLGENNELLPLEAVEPCSAHRRYFLPGS 602
 DB 612 SMRVPEGRTCSRLPLVTFEH-NGTVIEGQDGDNDNELLISRDLEPCTGNHRRYFKLGS 670
 QY 603 GYALFENYFVKWDAADAIQIASTVEVLNLTLEDEILPLSVYTKKEELRDVGVLDAEV 662
 DB 671 GYVYEDYNYRMEVP--ETISTRVTLNLTLEDEILPLSVYTKKEELRDVGVLDAEV 728
 QY 663 ARNQLHELFVDINKVLEVDN-----YAGLQEF 693
 DB 729 QRNQLHALKEFYDVRVVKVDHNVLLRGIANFQFGLGVDG 769

RESULT 10

AAR27807

ID AAR27807 standard; Protein; 933 AA.

XX AAR27807;

XX 10-MAR-1993 (first entry)

XX Bovine herpes virus type 1 glycoprotein I.

XX BHV-1; vaccine; GI.

XX Bovine herpes virus type 1.

XX Key Location/Qualifiers

XX Peptide 1..505

XX /note= "gIb peptide"

XX Peptide 506..933

XX /note= "gIc peptide"

XX Region 767..828

XX /note= "putative transmembrane sequence"

XX US5151267-A.

XX 29-SEP-1992.

XX 15-JUL-1988; 88US-0219939.

XX 15-JUL-1988; 88US-0219939.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Babiuk L, Fitzpatrick D, van den Hurk S, Zamb T;

XX WPI: 1992-348566/42.

XX N-PSDB; AAQ29515.

XX Protecting bovine hosts from bovine herpes virus type 1 infection -
 using vaccine contg. polypeptide neutralising epitope of GI, GII or
 GIII BHV-1 glyco:protein

XX Disclosure; Fig 5; 48pp; English.

XX The sequence is that of the BHV-1 glycoprotein gI which comprises the
 gIa/gIb/gIc complex, the glycoprotein maintains the proper epitopes
 necessary to fully protect immunised animals from disease. It elicits
 antibodies that neutralise virus infectivity and induce complement-
 mediated cell lysis. It can be used in vaccine compns. which are
 more protective than killed virus and attenuated live-virus vaccines.
 XX See also AAR27808 and AAR27809.

SQ Sequence 933 AA;

Query Match 31.1%; Score 1746; DB 13; Length 933;

Best Local Similarity 49.8%; Pred. No. 1.2e-101;

Matches 341; Conservative 106; Mismatches 218; Indels 20; Gaps 6;

QY 18 GTNSSPSTQNTVSREVSSVQLSEESTFYLCPPPGVSTVIRLEPPKCPPEPKATEW 77

DB 97 GDAASPNDSTVRAALRLAQAAGENSREFFVCPSPSGATVVRALARPCEYGLGRNTE 156

QY 78 GIALFKENISPKFKVTLTKYKNIQTWTGTYRQITNRYTDRTPVSEIEITDLIDGK 137

DB 157 GIGVYKENIAPYTKAVIYKNIYVTVTTWAGSTAAITNQYTRDVPVCMGEITDLVDK 216

QY 138 GRCSSKARYLRNNVYFAFDORAGEKQVLLKPSKFNTPESRAWHHTTNETYVMSPI 197

DB 217 WRCLSKAEYLRGKRVAFDRDDDDPWEAPLKPRLSAPGVRGWHHTDDVYTALGSAGLYR 276

QY 198 TGTSVNCIVEEMDARSVEPYSPAMANGDIANISPFYGLSPPEAAAPMGYPQDNFKOLD 257

DB 277 TGTSVNCIVEEARSVPYDPSFALSTGDIYNSPFYGLR-EGAHREHTSYSPERQOIE 335

QY 258 SYFSMDLDRKRRKASLPVKRNLITSHFTVGVWDAKPTTRVCSTKWKVEYTEMRLATVNGR 317
 Db 336 GYIKRDMATGRKLEKPEVSRNEFRTQHTVTVAWDVPRKKNVCSLAKWREADEMLRDESRGN 395
 QY 318 YRFMARELSATFISNTTEFDNRIILGQCICKREAAAEIQIFRTKYNDSHVKGHYQYFL 377
 Db 396 FRETARLSATFVSDSHTFALQNVPLSDCVIEEAAAEVRYERYNGTHVLSGSLETYL 455
 QY 378 ALGGFIVAYQVLSLAHMYRLMRDNRDMDLNVNKHAIYKKNATSLSRRLRDR 437
 Db 456 ARGGFVAFRPMLSNELAKLYLQELARSNGTLEGLFAA----AAPKPGP---RRARRAAP 508
 QY 438 NAPNRKITL-----DDTTAKSTSSVQFAMQLFLYDHIOTHINDMFSRIATANCELO 490
 Db 509 SAPGPGAANGPAGDGDAGRVTTVSSAEFAALQFTYDHIQDHVNTMFSRLATSWCLQN 568
 QY 491 RELVLWHEGKIKINPSATATLGRVAAKMLGDVAASVSCCTAIDAEVTLQNSMRVITST 550
 Db 569 KERALLWAEAAKLNPSAAASALDRRAAARMGLDAMAVTYCHELGEGRVFIENSMR--APG 526
 QY 551 NTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENY 610
 Db 627 GVCYSRPPVPSFAFGNESEVEGOLGEDNELLPGRELVEPCTANHKKRYFRFGADYVYENY 686
 QY 611 NFVKMVDAAADIOIATSFVELNLTLLDREILPLSVYTKKEELRDVGVLDAEVARRNOLHE 670
 Db 687 AYVRRVPLAELEVISFTVDNLNLTVDRELFPLEVYTRAEALADTGLLDYSEIQRRNOLHE 746
 QY 671 LKFDYINKVIEVDNTYA---GLQEF 692
 Db 747 LRFYDIDRVVKTDGNMAMIRGLANF 771

RESULT 11

AAR41343
 ID AAR41343 standard; Protein; 933 AA.

XX AC AAR41343;

DT 18-FEB-1994 (first entry)

DE Bovine herpesvirus type 1 gI glycoprotein.

KW BHV-1; Bovine herpesvirus; vaccine; neutralising epitope;
 KW glycoprotein; coat protein; gI; herpes; virus.

XX OS Bovine herpesvirus.

XX FH Key Location/Qualifiers

FT Cleavage-site 505 /note= gIa precursor protein cleaved to give gIb
 and gIc

FT Region 767..828 /label= Putative transmembrane region.

XX CA2057387-A.

XX PD 12-JUN-1993.

XX PF 11-DEC-1991; 91CA-2057387.

XX PR 11-DEC-1991; 91CA-2057387.

XX (VETE-) VETERINARY INFECTIOUS DISEASE.

XX PA Babluk L, Fitzpatrick D, Van DER HURK S, Zamb T;

XX DR WPI; 1993-273267/35.

XX DR P-PSDB; AAR41343.

XX PT Recombinant bovine herpes type 1 proteins gI, gII and gIV - and
 PT DNA encoding them, for use in vaccines against e.g. shipping
 fever

XX PS

XX Claim 2; Figure 5; 154pp; English.

CC The bovine herpesvirus gI glycoprotein (BHV-1) is used in the
 CC construction of vectors which are subsequently used to transform
 CC host cells. gI produced by these cells can then be used to produce
 CC subunit vaccines comprising one or more neutralising epitopes of the
 CC gI glycoprotein. The subunit vaccines are used to protect cattle
 CC from disease, especially shipping fever which is a complex syndrome
 CC which often includes infection by BHV-1. They are substantially
 CC more protective than previous killed or live attenuated virus
 CC vaccines.

XX SQ Sequence 933 AA;

Query Match 31.1%; Score 1746; DB 14; Length 933;
 Best Local Similarity 49.8%; Pred. No. 1.2e-101;
 Matches 341; Conservative 106; Mismatches 218; Indels 20; Gaps 6;

QY 18 GTSNSPSTQNTVSREVSVSSQVLSSEESTFYLCPPVGVSTVIRLEPPRKPPEPRKATEWGE 77

Db 97 GDDAASPDSNDVRAALRLAQAGENSREFVCPPPSGATVVRAPARPCPEYGLGRNYTE 156

QY 78 GIALPKENISPVYKFKVLYKNIQTTTGTCTYQITNRYTDRTPVSTEEITDLIDGR 137

Db 157 GIGVITKENIAPTTFKAYIYIRNVITTWAGSTYAAITNQYDTRVPVCGEITDLVDKR 216

QY 138 GRCSKARYLDRNNVYVEAFDRAGEKQVLLKPSKFTPESSRAWHHTTNETYTVMGSPWIYR 197

Db 217 WRCLSKAEVLRSGRKVVAFDRDDDPWEAPLKPARLSAPGVGRWHTTDDVYTALGSAGLYR 276

QY 198 TGTSVNCIVEEMDARSVPFYSYFAMANGDIANISPFYGLSPPEAAAPGPGYPODNFKOLD 257

Db 277 TGTSVNCIVEEVEARSVPYDSFALSTGDIYMSPFYGLR-EGAHREHTSYSPERFQIE 335

QY 258 SYFSMDLDRKRRKASLPVKRNLITSHFTVGVWDAKPTTRVCSTKWKVEYTEMRLATVNGR 317

Db 336 GYIKRDMATGRKLEKPEVSRNEFRTQHTVTVAWDVPRKKNVCSLAKWREADEMLRDESRGN 395

QY 318 YRFMARELSATFISNTTEFDNRIILGQCICKREAAAEIQIFRTKYNDSHVKGHYQYFL 377

Db 396 FRETARLSATFVSDSHTFALQNVPLSDCVIEEAAAEVRYERYNGTHVLSGSLETYL 455

QY 378 ALGGFIVAYQVLSLAHMYRLMRDNRDMDLNVNKHAIYKKNATSLSRRLRDR 437

Db 456 ARGGFVAFRPMLSNELAKLYLQELARSNGTLEGLFAA----AAPKPGP---RRARRAAP 508

QY 438 NAPNRKITL-----DDTTAKSTSSVQFAMQLFLYDHIOTHINDMFSRIATANCELO 490

Db 509 SAPGPGAANGPAGDGDAGRVTTVSSAEFAALQFTYDHIQDHVNTMFSRLATSWCLQN 568

QY 491 RELVLWHEGKIKINPSATATLGRVAAKMLGDVAASVSCCTAIDAEVTLQNSMRVITST 550

Db 569 KERALLWAEAAKLNPSAAASALDRRAAARMGLDAMAVTYCHELGEGRVFIENSMR--APG 626

QY 551 NTCYSRPLVFSYGENOGNIQOLGENNELPTLEAVEPCSANHRRYFLFGSGYALFENY 610

Db 627 GVCYSRPPVPSFAFGNESEVEGOLGEDNELLPGRELVEPCTANHKKRYFRFGADYVYENY 686

QY 611 NFVKMVDAAADIOIATSFVELNLTLLDREILPLSVYTKKEELRDVGVLDAEVARRNOLHE 670

Db 687 AYVRRVPLAELEVISFTVDNLNLTVDRELFPLEVYTRAEALADTGLLDYSEIQRRNOLHE 746

QY 671 LKFDYINKVIEVDNTYA---GLQEF 692

Db 747 LRFYDIDRVVKTDGNMAMIRGLANF 771

RESULT 12

AAY32470

ID AAY32470 standard; Protein; 943 AA.

XX AC AAY32470;

XX 27-MAR-2000 (first entry)
 XX DNA encoding feline herpesvirus-1 glycoprotein B.
 XX FHV; glycoprotein B; raccoon poxvirus; vaccine; cat;
 XX feline viral rhinotracheitis.

Feline herpesvirus type 1.

XX Key Location/Qualifiers
 XX Peptide 1..66
 XX /note= "signal peptide"
 XX Domain 67..757
 XX /note= "extracellular domain"
 XX Domain 758..827
 XX /note= "transmembrane domain"
 XX Domain 828..911
 XX /note= "cytoplasmic domain"
 XX Modified-site 152..154
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 261..263
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 364..366
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 406..408
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 526..528
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 610..612
 XX /note= "Asn is N-glycosylated"
 XX Modified-site 643..645
 XX /note= "Asn is N-glycosylated"

US6010703-A.

04-JAN-2000.

14-AUG-1997; 97US-0911321.

26-JUL-1993; 93US-0096183.

(UNMS) UNIV MICHIGAN STATE.

Spatz SJ, Maes RK;

WPI; 2000-115851/10.

N-PSDB; AAZ35417.

New recombinant raccoon poxvirus expressing glycoprotein from feline herpes virus, used in vaccines to protect against feline viral rhinotracheitis

Claim 3; Column 31-35; 57pp; English.

This sequence represents glycoprotein B (gB) of feline herpesvirus type 1 (FHV-1) C-27 (ATCC VR 636). The sequence was deduced from a 9B coding region identified in a SalI subfragment (see AAZ35417) of the FHV-1 genome. The invention provides a recombinant raccoon poxvirus which expresses a gene that encodes the FHV gp glycoprotein precursor (see AAY32673) and is inserted into the thymidine kinase (TK) gene of the poxvirus. The recombinant poxvirus optionally also contains a gene encoding the FHV-1 gB precursor inserted, in addition to the gp gene, into the TK gene of the poxvirus. A live vaccine containing the recombinant raccoon poxvirus is also claimed. The vaccine is used to protect cats against feline viral rhinotracheitis caused by FHV-1. It is safe, stable and more immunogenic than known modified live vaccines, providing better protection against development of clinical signs and viral shedding.

Sequence 943 AA;

AA50113

ID AAM50113 standard; protein; 943 AA.

XX

Query Match 31.08; Score 1743; DB 21; Length 943;
 Best Local Similarity 40.3%; Pred. No. 2e-101;
 Matches 377; Conservative 161; Mismatches 267; Indels 130; Gaps 21;
 QY 3 YFRNCIF-----FLIVLY-----GTSSP 23
 DB 22 YFRCFFPGLGTAATGSRHNGSSGLTRARVSVFIWVFLVGRPRVGGSGSSEQ 81
 QY 24 STQNTSREV-----VSSVQSEEE-----STFYLCPPPVGVSTVIRLEPP 63
 DB 82 PRRTVATPEGVHQNOLQIPPCRYEEALRASQIEANGSPSTFYMCPPSPSVVRLEPP 141
 QY 64 RKPCEPRKATWEGGIALFKENISPKYFKVLYKNIQTTWTGTYRITRYDRT 123
 DB 142 RACPDYKLGKNTFEGIAVIFKENTAPIYKFKANIYKNIIMTVWSSGSYATVTRYDTRV 201
 QY 124 PVSEETDLDGKRCSSKARYLRNVVYFAFDRDAGEKOVLLKPSKFNTPESRAWHTT 183
 DB 202 PVKQETDLDLRGMCLSKADYVRNNYQFAFDRDDEPRELPLKPSSTLSRVRGWH-T 260
 QY 184 NETYVWSPWIYRTGTSVNCIVEEMDARSYFYSIFAMANGDIANISPFYGLSPPEAAA 243
 DB 261 NETYTKIVLLDFHSGTSVNCIVEEDARSYVPYDSFAISGTDVHMSPPFGLR-DGAHV 319
 QY 244 EPMGYPODNPKQLDSYFMDLDRKASLPYKRNFLITSHFTVGDWHPAKTRVCSMTKW 303
 DB 320 EHTSYSSDRFQOIEGYPIDLDTDTGA-PVSRNLETPTPHVTVAWNTPKSGRYCTLAKW 378
 QY 304 KEVTEMLRATVNGRYRFMARELSATFISNTFDPNRIILGQCICKRAEAAIEQIFRTKY 363
 DB 379 REIDEMPLMNI-GSYRTAKTISATFISNTSQFENRIRLGDCATKEAAEIDRIYKSKY 437
 QY 364 NDHVHVGHVQYFLALGGFIVAYOPVLSKSLAHMYLRELMRDNRDTEMLDVLNKNHAIYK 423
 DB 438 SKTHIQGTLETYLARGGFLIAPRPMISNELAKLYINELARSNRT--VVOL-----SALLN 491
 QY 424 KNATSLSLRREDIRNAPNRKI-----TLDDTTAIKTSSVQFAMQLFLYDHIQTH 473
 DB 492 PSGETVQTRKRSVPNSQNHRSRRRTIEGGIETVNNASLLKTTSSVEFAMLFADYIQAH 551
 QY 474 INDMFSRIATAWCELORELVLWHEGINKINPISATASATGLRRVAAKMLGDVAAVSSCTAI 533
 DB 552 VNEMLSIATATWCTLQNRHVLWTETLKLNPGGVVSMALERRVSARLLGDGAVAVTQCVNI 611
 QY 534 DAESVTLQNSMRVITSTNTCYSRPLVLFSGYENOGNTQGLGENNELLPTLEAVEPCSAN 593
 DB 612 SSGHYIQNSMRVTGSSSTTCYSRPLVSFRALNDSEYIEGQLGENNELLVERKLEPCTVN 671
 QY 594 HRRYFLFGSGYALFENYFVKWDAADIQIASTFVELNLTLEDREILPLSVYTKLELRD 653
 DB 672 NKRYFKGADYVYFEDYAVYKRVPLSEIELISAVY-INKSTLLEDFLH-SSYTRAELED 729
 QY 654 VGLDYAEVARRNQLHELFYDINKVIEVDTPN-----YAGLQFEGC----- 694
 DB 730 TGPPDYSEIORRNLHALKFYDIDSIVRVDNVLIMRGMANFFQGLGDVGDFGKVVGLGA 789
 QY 695 -----MSITKDPANPNNGQTQLEARMELTDLIINA-----KAM---TLAS 731
 DB 790 ASAVISTVSGVSSFLNPPFGALAVGLLILAGVIAAFAYRIYISLRANPKMALPYVTRN 849
 QY 732 LDYAKTEASLSAYSEATVNNLNATLEQLKAK-TNLESAINQANTDRTTDFEDNHPN 790
 DB 850 LKQAKSPASTAGGSDPGVDDEEKLQMQAREMIKYMSIVLSAMEQ-QEHKAMKNGKPA 908
 QY 791 LVEAY----KALKT--TLEQRATNLEGLSSSTAYNQI 820
 DB 909 ILTSHLTNMLRRRGPVKYQRLNNLDSGDDTETNLV 943

RESULT 13
 AAM50113
 ID AAM50113 standard; protein; 943 AA.
 XX

DR WPI; 1996-105220/11.
XX N-PSDB; AAT16474.
PT Detection of herpes B virus by PCR amplification of sample DNA - to
PT detect a specific herpes simian monkey B virus DNA segment.
XX
PS Claim 1; Column 19-24; 22pp; English.
XX
CC The herpes simian monkey B virus proteins, such as the gB
CC glycoprotein (UL27), have immense potential use in the development
CC of serological immunoassays, which can specifically detect virus
CC antigens and/or antibodies to B virus. One approach is to
CC synthesize peptides which, based on the properties of the predicted
CC protein sequence, are likely to be immunologically active. Such
CC peptides can be used as substrate antigens in immunoassays to detect
CC serum antibodies which recognize this specific peptide sequence.
CC Synthetic peptides may also be used to produce antibodies against
CC specific regions of the gB glycoprotein which are unique to one
CC virus. These can then be used to develop virus-specific
CC immunoassays for differentiation of B virus from other primate
CC alpha-herpes virus and for identification of antibodies directed
CC against B virus in primate serum samples.
XX
SQ Sequence 891 AA;
Query Match 31.0%; Score 1739.5; DB 17; Length 891;
Best Local Similarity 49.0%; Pred. No. 3e-101;
Matches 332; Conservative 103; Mismatches 216; Indels 27; Gaps 3;
Qy 18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPPGVSTVIRLEPPKCPKPRKATEWGE 77
Db 72 GTNASVEAGHATLRENLDKALGDATFYVCPPTGATVQVQPPRCPCPRAPHGQNYTE 131
Qy 78 GIATLFENISPYKFKVLYKNIQTWTGTYRQITNRYTDRTPVSEIEITDLIDGK 137
Db 132 GIAVIFKENIAPYKFKATMYKQVTSQVWFGHYSQFMGIFEDRAPVPFEEVDKINAR 191
Qy 138 GRCSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPESEAWHTTNETYTVWGSPIYR 197
Db 192 GVCSTAKYVRNNMESTAFHRDDDESDMKLPAKAAATRTSGWHTTDLKYNPSRIEAFHR 251
Qy 198 TGTSVNCLVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMGYPQDNFKOLD 257
Db 252 YGTTVNCIVEARSVPYDFEVLATGDFVYMSPFYGYR-DGAHAETHAYAAADRFRQVD 310
Qy 258 SYFSMDLDRKASPLKYNRLIISHTYVGDWAPKTRVCSMTKWEVTEMLATVNGR 317
Db 311 GYERDLSTGRSTPATRNLLTTPKFTVGDWAPKRPVCTLLTKWQEVDEMLRAEYGPS 370
Qy 318 YRFMARELSATFISNTTFDNPRIILGQIKREAEAEQIFRTKYNDSHVKVGHVOYFL 377
Db 371 FRFSSALSTFTTNRTEYALSVDLGDVCGREAREAVDRIFLRVNGTHVKVQVQYLL 430
Qy 378 ALGGFIVAYQVPLSKSLAHMYLRELNRDNRTEMLDLNKNKHAIYKKNATSLSLRRDIR 437
Db 431 ATAGFLIAYQPLLSNGLVLYVRELLREQGRP-----GDAA 467
Qy 438 NAPNRKITLDTTATKSTSSVQFAMQLQFLYDHIQTHINDMFSRTATANCELORELVLWH 497
Db 468 ATPASDPDPPOVERIKTSSVEFARLQFTYDHIQRHNDMLGRILAIWCELQNELTLWN 527
Qy 498 EGKINPSATASATIGRAVAAKMGDVAASVSCATAIDAESVTLONSNRVITSTNCTSRP 557
Db 528 EARLKNPAIASATVGRVSRMGLDVMNAVSTCVPTVPDVMNSMRKVPAPGTCISRP 587
Qy 558 LVLFSGYNOGNIQOLGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENYFVKMVD 617
Db 588 LVSFYRGGPLVEGQLEDNEIRLERDALEPCTVGHRRYFTFGAGYVFEYDAYSHQLG 647
Qy 618 AADTQIASTFFVELNLTLEDEILPLSVYTKELRDVGLDYAEVARNQHLKFKVDIN 677
Db 648 RADVTVTSTFINLNLTMLEDEHFVPLEVYTRQIEIKSGLLDYTEVQRNQLHALRFADID 707

Qy 678 KVIEVDTN---YAGLQEF 692
Db 708 TVIKADAHAPLFAFLYSF 725
RESULT 15
AAW70293
ID AAW70293 standard; Protein; 891 AA.
XX
AC AAW70293;
XX
DT 06-NOV-1998 (first entry)
XX
DE Simian herpesvirus B gB glycoprotein sequence (UL27).
XX
KW Simian herpesvirus B gB glycoprotein; UL27; ICP protein; UL28;
KW differential diagnostic test; immunoassay; antibody.
XX
OS Simian herpesvirus B.
XX
FN US5767265-A.
XX
PD 16-JUN-1998.
XX
PF 10-OCT-1995; 95US-0541878.
XX
PR 01-APR-1993; 93US-0042747.
PR 10-OCT-1995; 95US-0541878.
XX
PA (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.
PI Black D, Eberle R, Hilliard J, Scinicariello F;
XX
DR WPI; 1998-361791/31.
DR N-PSDB; AAV33167.
XX
PT Monkey herpes B virus DNA - coding for gB glycoproteins and
PT polypeptides
XX
PS Claim 2; Columns 19-24; 22pp; English.
XX
CC The invention provides a simian herpesvirus B DNA sequence coding
CC for the present gB glycoprotein (UL27) and a portion of an ICP
CC 18.5 kDa protein (UL28; AAW70294). The invention uses these DNA and
CC protein sequences as a basis for the development of differential
CC diagnostic tests for the rapid identification of simian herpesvirus B
CC cases. Therefore, the virus can be detected by detecting the DNA
CC sequence and knowledge of the amino acid sequence will help in the
CC design of DNA probes and of peptides for use in immunoassays and for
CC antibody production.
XX
SQ Sequence 891 AA;
Query Match 31.0%; Score 1739.5; DB 19; Length 891;
Best Local Similarity 49.0%; Pred. No. 3e-101;
Matches 332; Conservative 103; Mismatches 216; Indels 27; Gaps 3;
Qy 18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPPGVSTVIRLEPPKCPKPRKATEWGE 77
Db 72 GTNASVEAGHATLRENLDKALGDATFYVCPPTGATVQVQPPRCPCPRAPHGQNYTE 131
Qy 78 GIATLFENISPYKFKVLYKNIQTWTGTYRQITNRYTDRTPVSEIEITDLIDGK 137
Db 132 GIAVIFKENIAPYKFKATMYKQVTSQVWFGHYSQFMGIFEDRAPVPFEEVDKINAR 191
Qy 138 GRCSKARYLRNNVYVEAFDRDAGEKQVLLKPKSFNTPESEAWHTTNETYTVWGSPIYR 197
Db 192 GVCSTAKYVRNNMESTAFHRDDDESDMKLPAKAAATRTSGWHTTDLKYNPSRIEAFHR 251
Qy 198 TGTSVNCLVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMGYPQDNFKOLD 257
Db 252 YGTTVNCIVEARSVPYDFEVLATGDFVYMSPFYGYR-DGAHAETHAYAAADRFRQVD 310


```
QY 258 SYFSDLDKRRKASLPVKRNLTSHTFTVGWQWAPKTRVCSMTKKWKEVTEMLRATVNGR 317
Db 311 GYERDLSTGRRASTPATRNLLTTPKFTVGWQWAPRPSVCTLTWKQEVDEMLRAEYGPS 370
QY 318 YRFMARELSATFISNTTEFDPNRIILGQCICKREAAAEIQIFRTKYNDSHVKVGHVQYFL 377
Db 371 FRESSALSTTTTNTTEALSRLVDGDCVGREAREAVDRIFLRRYNGTHVKVGVQYYL 430
QY 378 ALGGFIVAYQPVLSKSLAHMYLRELMRDNRDTEMLDLVNNKHAHYKKNATSLSLRRDIR 437
Db 431 ATAGFLIAYQPLLSNGLVLYRELLREQGRP-----GDA 467
QY 438 NAPNRKITLDDTTAIAKSTSSVOFAMLOELYDHIQTHINDMESRIATAWCELONRELVLWH 497
Db 468 ATPKPSADPPVERIKITTSVEFARLQFTDHIQRHVNDMLGRIATAWCELONHELTWN 527
QY 498 EGIKINPSATASATLGRRAAKMLGDVAAVSSCTAIDAESVTIQNSMRVITSTNTCYSRP 557
Db 528 EARKLPNAIASATVGRVRSARMLGDVMAVSTCVPTPDNVIMQNSMRVPAREGTCYSRP 587
QY 558 LVLFSGENOGNIQOLGENNELLPTLEAVEPCSANRRYFLFGSGYALFENYFVKMVD 617
Db 588 LVSFRYEEGGPLVEGOLGEDNEIRLERDALEPCTVGHRRYFTFGAGYVYFEDYAYSHQLG 647
QY 618 AADIQIATAFVELNLTLLEDREITPLSVYTKKELRDVGVLDAEVARRNOLHELKEYDIN 677
Db 648 RADVTTVSTFINLNLTMLEDHEFVPLEVYTRQIKDSGLLDYTEVQRRNOLHALRFADID 707
QY 678 KVIEDTN---YAGLOEF 692
Db 708 TVIKADAHAPLFLAGLYSF 725
```

Search completed: October 8, 2003, 17:02:51
Job time : 91.2345 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1736	30.6	904	1	PCR-US03-11231-18	Sequence 18, Appl
2	681	12.0	854	5	US-09-350-841A-1589	Sequence 1589, Ap
3	181.5	3.2	770	5	US-09-897-516A-4453	Sequence 4453, Ap
4	158	2.8	807	5	US-09-820-843B-108	Sequence 108, Appl
5	151.5	2.7	2653	7	US-60-490-890-1432	Sequence 1432, Ap
6	149	2.6	3418	1	PCR-US03-24394-1	Sequence 1, Appl1
7	147	2.6	1261	1	PCR-US03-36491-192	Sequence 192, Appl
8	147	2.6	1261	6	US-10-648-593-192	Sequence 192, Appl
9	146.5	2.6	523	6	US-10-425-114A-58966	Sequence 58966, A
10	146.5	2.6	1033	5	US-09-820-843B-75	Sequence 75, Appl
11	145	2.6	1323	1	PCR-US03-19153-312	Sequence 312, Appl


```

Db      786 ValAlaAlaPheAlaPheArgTyrValLeuGlnLeuGlnArgAsnProMetLysAla 805
QY      2275 -----ACATTAGAACAACTAAATAATGGCTAAACATAATTAGAAATCAGCCATC 2322
Db      806 LeuTyrProLeuThrLysGluLeuLysThrSerAspProGlyGlyValGlyGlyGlu 825
QY      2323 ACCAAGCTAATACGATAAAGGACTTTTGTATATGAATGAACACCCAAATTTAGTTGAAGCA 2382
Db      826 GlyGluGluGlyAlaGluGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
QY      2383 -----TACAAACACTAAACAAACCCACTTTAGAA-----CAACGTGCTACT 2421
Db      844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
QY      2422 AACCTTGAAGTTGTGTCATCACTGCTTATAATCAAAATTCGCAATAATTTAGTGATCA 2481
Db      864 ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys 883
QY      2482 TACAATAAAGCT 2493
Db      884 ArgAsnLysAla 887

RESULT 2
US-09-350-841A-1589
; Sequence 1589, Application US/09350841A
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1589
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-350-841A-1589

Alignment Scores:
Pred. No.:      4,94e-40      Length:      854
Score:          681.00      Matches:      211
Percent Similarity: 43.76%      Conservative: 122
Best Local Similarity: 27.73%      Mismatches: 338
Query Match:      12.02%      Indels:      90
DB:              5          Gaps:         21

US-09-147-052-3 (1-3261) x US-09-350-841A-1589 (1-854)
QY      16 CGGAATTCGATATTTTCCTTATAGTATTCTATATGGT----- 54
Db      3 ArgArgValLeuSerValValLeuLeuAlaLeuAlaLeuAlaCysArgLeuGlyAla 22
QY      55 ----ACGAATCATCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTC 111
Db      23 GlnThrProGluGlnProAlaProAlaThrThrValGlnProThrAlaThrArgGln 42
QY      112 CAGTTGTCGTGAGGAGAGCTACGTTTATCTTTGTCCTCCCAAGTGGTTCACACCGTG 171
Db      43 GlnThrSer-----PheProPheArgValCysGluLeuSerSerHisGlyAspLeu 59
QY      172 ATCCGTCGTAGACCGCGGAAATGTCCTCCGAACCTAGAAAGCCACCGAGTGGGTGAA 231
Db      60 PheArgPheSerSerAspIleGlnCysProSerPheGlyThrArgGluAsnHisThrGlu 79
QY      232 GGAATCGCATATATTAAAGAGATATCATCTCCATATAAATTTAAAGTGAGCTTTAT 291
Db      80 GlyLeuLeuMetValPheLysAspAsnIleIleProTyrSerPheLysValArgSerTyr 99
QY      292 TATAAAATATCATTCAGACGACGACATCGACGGGACGACATATAGACAGATCACTAAT 351
Db      100 ThrLysIleValThrAsnIleLeuIleTyrAsnGlyTyrTyrAlaAspSerValThrAsn 119

```

```

QY      352 CGATATACAGATAGGAGCGCCGTTTCCATTGAAGAGATCAGCGATTAATCGACGGCAA 411
Db      120 ArgHisGluGluLysPheSerValAspSerTyrGlu---ThrAspGlnMetAspThrIle 138
QY      412 GGAAGATGCTCATCTCTAAACAAGATACCTTTAGAAAC-----AATGTATATGTTGAA 462
Db      139 TyrGlnCysTyrAsnAlaValLysMetThrLysAspGlyLeuThrArgValTyrVal--- 157
QY      463 GCGTTTGCAGGATCGCGGAGAAAAACAAGTACTTCTAAACCAATCAAAATTCACACG 522
Db      158 -----AspArgaspGlyValAsnIleThrValAsnLeuLysProThrGlyGlyLeuAla 175
QY      523 CCCGAATCTAGGCGATGGCACAGCACTAATAGACAGCTATACCGTGGGGA---TCACCA 579
Db      176 AsnGlyValArgArgTyrAlaSerGlnThrGluLeuTyrAspAlaProGlyTyrLeuIle 195
QY      580 TGGATATATCGAAGGGAACCTCCGTCATTTGTATAGTAGAGAAATGATGCCGCTCT 639
Db      196 TrpThrTyrArgThrArgThrValAsnCysLeuIleThrAspMetMetAlaLysSer 215
QY      640 GTGTTTCCGTATTTCATATTTGCAATGGCCAAATGGCGACATCGCGAACAATATCTCATT 699
Db      216 AsnSerProPheAspPheValThrThrGlyGlnThrValGluMetSerProPhe 235
QY      700 TATGTTCTATCCCAACAGAGCGTCCGCGAGAACCCATGGGATATCCCGAGTAATATTC 759
Db      236 TyrAspGlyLysAsnLysGluThrPheHisGlu-----ArgAlaAspSerPhe 251
QY      760 AAACAATAAGTAGATGATTTTCAATGGATTGGACAAGCGTGGAAAGCAAGCCCTTCCA 819
Db      252 HisValArgThrAsnTyrLysIleValAspTyrAsnArgGlyThrAsnProGlnGly 271
QY      820 GTCAAGCGTAACCTTCTCATCATCACACTTCACAGTTGGGTGGGACGTTGGCTCCAAA 879
Db      272 GluArgArgAlaPheLeuAspLysGlyThrThrLeuSerTrpLysLeuGluAsnArg 291
QY      880 ACTACTCGGTATGTTCAATGACTAAAGTGGAAAGAGGTGACTGAAATGTTGCGTGCACA 939
Db      292 ThrAla---TyrCysProLeuGlnHisTrpGlnThrPheAspSerThrIleAlaThrGlu 310
QY      940 GTTAATGGGAGATACAGATTATTATGGCCCGTGAACCTTCGCGAAGTTTATCAGTATACG 999
Db      311 ThrGlyLysSerIleHisPheValThrAspGluGlyThrSerSerPheValThrAsnThr 330
QY      1000 ACTGAGTTTGTATCCAAATCGCATCATATTAGGA-----CAATGT 1038
Db      331 Thr-----ValGlyIleGluLeuProAspAlaPheLysCys 342
QY      1039 ATTAACGGGAGGACAGACAGCAATCGACAGATATTAGGACAAATATATGACAGT 1098
Db      343 IleGlu-----GluGlnVal-----AsnLysThr 350
QY      1099 CAGGTCAAGTTGGACATGTACNA-----TAT 1125
Db      351 HisGluLysTyrGluAlaValGlnAspArgTyrThrLysGlyGlnGluAlaIleThrTyr 370
QY      1126 TTCCTTGGCTCTCGGGGATTTATGTAGCATATCAGCTGTCTTATCCAAATCCCGGCT 1185
Db      371 PheIleThrSerGlyGlyLeuLeuAlaTrpLeuProLeuThrProArgSerLeuAla 390
QY      1186 CATATG---TACCTCAGAAATGTGAGAGACACAGGACCGCATGCTGCACCTG 1242
Db      391 ThrValLysAsnLeuThrGluLeuThrThrProThrSerSerProProSerProSer 410
QY      1243 GTAACAATAAGCATGCAATTTATAGAAAAATGCTACCTCATTTGTCAGATTGCGGCGA 1302
Db      411 ProProAlaProSerAla-----AlaArgGlySerThrProAlaAlaValLeuArgArg 428
QY      1303 GATATTCGAATGCCCAATAGAAAAATAACATTAGACAGACACCACTATTAAATCG 1362
Db      429 ArgArgArgAspAlaGlyAsnAlaThrThrProValProThrAlaProGlyLysSer 448

```

```

QY 1363 ACATCGCTGTTCAA-----TTGCCCATGCTCCAAATTTCTTTATGATCATATACAA 1413
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1414 ACCCATATTATGATATGTTAGTAGGATGCCACAGCTGGTCCGAATTCGCAATAGA 1473
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1474 GAACTGTTTATGCGCAGGAGGATAAAGATTAACTAGCGCTACAGCGAGGCAACA 1533
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1489 GlnAsnMetValLeuArgGluLeuThrLysIleAsnProThrValMetSerSerIle 508
QY 1534 TTGAGAGGAGAGGCTGCAAGATGTTGGGGATGCTGCTGCTATCGAGCTGCAC 1593
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1594 GCTATAGATGGGAATCCGTCACATTTGCAAAATCTATGGAGTTATCACATCCAC 1653
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1654 ACATGTTATACCGCAGCATGTTCTATTTTCATATGGAGAAACCAAGGAACATACAG 1713
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1714 GGCAACTCGGTGAAACACAGCTGCTTCCCAACGCTAGAGGCTGTAGAGCCATGCTCG 1773
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1774 GCTAATCATCGTAGATATTTCTGTTGGATCCGCTGTTATGCTTTATTTGAAACATATA 1833
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1834 TTTGTTAAGATGCTAGACGCTGCCGATATACAGATTGCTAGCACATTTGCTGACCTTAAT 1893
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1894 CTAACCTGCTAGAGATCGGGAATTTTGCTTTATCCGTTTACACAAAGAGAGTTG 1953
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1954 CGTGATGTTGTTGATTTAGTATGACAGAGTACCTGCCCAATCAATCAATGAACTT 2013
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 2014 AAATTTATGACATAACAAAGTATAGAGTGGATGATCAATACGGGGCTCGAGAA 2073
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 2074 TTCGCTGTATGCTATTACTATAAAGATGCAACCCAAATAATAGCCAAACCAATTA 2133
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 2134 GAAGCAGCGGAATGGAGTTAACAGATCTAATCAATGCTAAGAGCATGACATTAAGCTTCA 2193
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 2194 CTA 2196
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

RESULT 3

```

US-09-897-516A-4453
; Sequence 4453, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

```

; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 4453
; LENGTH: 770
; TYPE: PRN
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-4453

Alignment Scores:
Pred. No.: 3.3e-05 Length: 770
Score: 181.50 Matches: 180
Percent Similarity: 36.23% Conservative: 141
Best Local Similarity: 20.32% Mismatches: 342
Query Match: 3.20% Indels: 223
DB: 5 Gaps: 45

US-09-147-052-3 (1-3261) x US-09-897-516A-4453 (1-770)
QY 769 GATAGCTATATTTTCAATG-----GATTGGACAAAGCGTGAAGCAAGCCTTCCAGCT 822
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 823 AAGCGTAACTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 882
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 883 ACTCGTGTATGTTCAATGACTAAGTGGAAAGAGGTGACTGAAATGTTGCGTGCAACAGTT 942
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 943 AATGGGAGATACAGATTATGCGCCGTGAACCTTTTCGGCAACGTTTATCATGTAATACGACT 1002
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1003 GAGTTTCAATCCAAATCCATCATATTTAGGACAATGTATT---AAACCGAGGAGCAAGCA 1059
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1060 GCAATCAGAGATATTTAGGACAAATAATATACAGTACAGTCACGTCACGTCAGTCAGT 1116
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1117 -----GTACAATATTTCTTGGCTCTCGG 1140
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1141 -----GGATTATTGTAGCATATCATGCTGTTCTATCCAAATCCCTGGCT 1185
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 125 AspArgAsnLeuThrGlyTrpIleAla-----GluLeuIleSerGlyLeuAlaThr 141
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1186 CATATGACCTCAGAGAATTGATGAGAGAC-----AACAGACCGCATGAG 1230
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 142 ArgSerTyrValArgGlyLysIleArgAspHisAlaGluSerArgAsnHisProAspAla 161
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1231 ATGCTGACCTGTTAAACAAATGCAATGCAATTTTAAAGAAAAATGCTACCTCATTTGCA 1290
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 162 ThrLeu-----LysAsnLysGlyPheValIleLeuSerAsnAlaValAsp----- 176
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1291 CGATTGCGGCGAGATATTCGAAATGCACCAATAGAAAATAACATATAGACGACACCACA 1350
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 177 -----SerAsnSerGluAlaHisAlaAlaThrSerLys 187
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1351 GCTATTAAATCGACATCGTGTTCATATTCGCCATGCTCCCAATTTCTTTATGATCATATA 1410
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 188 AlaValLysSerThrHisGlu-----LeuAlaGlnLysAlaTyr---AsnLeu 202
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1411 CAAACCCATATTAATGATATGTTTAGTAGGATTGCCACAGCTTGGTGC----- 1458
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 203 AlaAspProValAsnAlaMet-----LysLysTrpValProLeuThrArg 217
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```



```

; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1432
; LENGTH: 2663
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-490-890-1432

```

Alignment Scores:

```

Pred. No.: 0 00494 Length: 2663
Score: 151.50 Matches: 129
Percent Similarity: 36.97% Conservative: 108
Best Local Similarity: 20.12% Mismatches: 233
Query Match: 2.67% Indels: 171
DB: 7 Gaps: 29

```

US-09-147-052-3 (1-3261) x US-60-490-890-1432 (1-2663)

```

Qy 970 GAACCTTTGGCAGCGTTTATCAGTAATACG---ACTGAGTTTGTGATCCAAATCGCATATA 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1313 GluGlnSerThrThrLysAspSerThrThrLeuAlaArgIleGluMetGluArgLeuArg 1332

Qy 1027 TTAGGACAGTGTATAACGGCGGAGAGACAGACAGTATCGAGAGATATTAGGACAAA 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1333 LeuAsnGluLysPhe---GlnGluSerGlnGluIleLysSerLeuThrLysGluArg 1351

Qy 1087 TATAATGACAGTCACGCTCAAG-----CTTGGACATGACATATTCTTGGCT 1134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1352 AsnAsnLeuLysThrIleLysGluAlaLeuGluValLysHisAspGlnLeuLysGluHis 1371

Qy 1135 CTCGGGGGATTTATTAGCATATACAGCTGTTCTATCCAAATCCCTGGCTCATATGTAC 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1372 IleArgGluThrLeuAlaLysIleGlnGluSerGlnSerLysGlnGluGlnSerLeuAsn 1391

Qy 1195 CTCAGAGATTGATGAGACACAGCAGCGATGAGATGCTGACCTGGTAAACAATAAG 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1392 MetLysGlu-----LysAspAsnGluThrThrLysIleValSerGluMetGluGln--- 1408

Qy 1255 CATGCAATTATAGAAAATGCTACCTCATTTGTCAGATTCGCGGCGAGATATTCGAAT 1314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1409 -----PheLysProLysAspSerAlaLeuLeuArgIleGluIleGluMetLysGly 1425

Qy 1315 GCACCAATAGAAAATAACATTAGACGAC-----ACCAGCTATTAAATCGACATCG 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1426 LeuSerLysArgLeuGlnGluSerHisAspGluMetLysSerValAlaLysGluLysAsp 1445

Qy 1369 TCTGTTCAA-----TTGCGCATGCTCCATTTCTTTATGATCATATACAAACCATATT 1422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1446 AspLeuGlnArgLeuGlnGluValLeuGlnSerGluSerAspGlnLeuLysGluAsnIle 1465

Qy 1423 AATGATATGTTTACTAGG-----ATTGCCACAGCTTGTGTCGAA 1461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1466 LysGluIleValAlaLysHisLeuGluThrGluGluLeuLysValAlaHisCysCys 1485

Qy 1462 TTGCAAGTATGAGAACTGTTTATGCGCAGGATTAAGATTAATCTCCTAGCGCTACA 1521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1486 LeuLysGluGlnGluThrIleAsnGlnLeuArgValAsnLeuSerGluLysGluThr 1505

Qy 1522 GCGAGTGCACATTAGGAGAGAGTGGCTGCAAAAGATGTGGGGGATGCTGCTGTA 1581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1506 GluIleSerThrIleGlnLysGln----- 1513

Qy 1582 TCGAGCTGCACGTATAGATCGCGAATCCGTCACCTTGTGCAAAATTCATCGGAGTATC 1641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1514 -----LeuGluAlaIleAsnAspLysLeuGlnAsnLysIleGlnGluIle 1528

Qy 1642 ACATCCACTAATACATGTTATAGCGGACCATTTGTTCTATTATTCATATGGGAAACCA 1701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1529 -----TyrGluLys-----GluGluGln 1534

```

```

Qy 1702 GGAACATACAGGAGACAACTCGGTGAAACAAACAGAGTTGCTTCCAAACGCTAGAGCTGTA 1761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1535 LeuAsnIleLys---GlnIleSerGluValGlnGluAsnValAsnGluLeuLysGlnPhe 1553

Qy 1762 GAGCCATGCTCGGCTAATCATCATCGTAGATATTCTTGTGTTGGATCCGGTATGCTTTATTT 1821
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1554 Lys-----GluHisArg----- 1557

Qy 1822 GAAACATATAATTTTGAAGTAGTAGACGCTGCCGATATACAGATTGCTAGCACATTT 1881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1558 -----LysAlaLysAspSerAlaLeuGlnSerIleGluSerLysMet 1571

Qy 1882 GTCGAGCTTAATCTAACCCCTCTAGAGATCGGGAATTTTTCCTTTATTCGTTTACACA 1941
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1572 LeuGluLeuThrAsnArgLeuGlnGluSerGlnGluGluIleGlnIleMetIleLysGlu 1591

Qy 1942 AAAGAAGAGTTGCGTGATGTTGCTGATTTGATGATGATGATGATGATGATGATGATGATGAT 2001
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1592 LysGluGluMetLysArgVal-----GlnGluAlaLeuGlnIleGluArgAspGln 1608

Qy 2002 CTACATGAA----- 2010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1609 LeuLysGluAsnThrLysGluIleValAlaLysMetLysGluSerGlnGluLysGluTyr 1628

Qy 2011 -----CTTAAATTTTATGACATAAACAAGTAATAGAGTGGATACAAATTTACGCGGG 2064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1629 GlnPheLeuLysMetThrAlaValAsnGluThrGlnGluLysMetCysGluIleGluHis 1648

Qy 2065 CTCGAGGAATTCGCTGTATCTTACTAAAAAGATGCAACACCAAAATAAATGCGCCAA 2124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1649 LysGlu-----GlnPheGluThrGlnLysLeuAsnLeuGluAsn----- 1662

Qy 2125 ACCCAATTAGAACGACGCGAATGAGTTAAGATCTAATCAATGCTAAAGCGATGACA 2184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1663 -----IleGluThrGluAsnIleArgLeuGlnIleLeuHisGluAsnLeuGluGlu 1680

Qy 2185 TTAGCTTCATCA-----CAAGACTATGCCAAGATTGAAGCTAGTATTGATTCATCT 2232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 MetArgSerValThrLysGluArgAspAspLeuArgSerValGluGluThrLeuLys--- 1699

Qy 2233 GCTTATAGTGAAGTGAACAGTTAACAATACCTTAATGCTCAACATTA----- 2280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1700 -----ValGluArgAspGlnLeuLysGluAsnLeuArgGluThrIleThrArgAspLeu 1717

Qy 2281 -----GAACAACTAAAAATGCTAAAACTAATTTAGATGATCAGCCATCAACCAAGCT 2331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1718 GluLysGlnGluGluLeuLysIleValHisMetHisLeuLysGlu-----HisGlnGlu 1735

Qy 2332 AATACGGATAA----- 2349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1736 ThrIleAspLysLeuArgGlyIleValSerGluLysThrAsnGluIleSerAsnMetGln 1755

Qy 2350 TTTTCATATGACACCCCAATTTAGTTCAACATACAAAGCAGCTAAAAACCACTTTA--- 2406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1756 LysAspLeuGluHisSerAsn-----AspAlaLeuLysAlaGlnAspLeuLysIleGln 1773

Qy 2407 -----GAACAACTGCTACT-----AACCTT 2427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1774 GluGluLeuArgIleAlaHisMetHisLeuLysGluGlnGlnGluThrIleAspLysLeu 1793

Qy 2428 GAAGTTTGTCACTGCTTATTAATCAAAATTCGCAATAATTTAGTGTGCTATACATA 2487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1794 ArgGlyIleValSerGluLysThrAspLysLeuSerAsnMetGlnLysAspLeuGluAsn 1813

Qy 2488 AAACCTAGTACTTTAATAACTAAACACATAGATCCATTAATGGGGAACGCTTTTAGAT 2547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1814 SerAsnAlaLysLeuGlnGluLysIleGlnGlu-----LeuLys 1826

Qy 2548 TCTAATGAG-----ATTACTACAGCTAATAGATATTAATAATACGTTATCACTATT 2601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1827 AlaAsnGluHisGlnLeuLeuIleThrLeuLysLysAspValAsnGluThrGlnLysLysVal 1846

```

QY 2602 AATGAACAAAGACTAATGCTGATGATTATCTAATAGTTTATTATAAAGTGATTCAA 2661
 Db 1847 SerGluMetGluGln-----LeuLysLysGlnLys 1857
 QY 2662 AAT 2664
 Db 1858 Asp 1858
 RESULT 6
 PCT-US03-24394-1
 ; Sequence 1, Application PC/TUS0324394
 ; GENERAL INFORMATION:
 ; APPLICANT: The Wistar Institute
 ; APPLICANT: Ramin Shiekhatair
 ; TITLE OF INVENTION: METHODS FOR REGULATING BRCA1-BRCA2-CONTAINING COMPLEX ACTIVITY
 ; FILE REFERENCE: WSTR-00148
 ; CURRENT APPLICATION NUMBER: PCT/US03/24394
 ; CURRENT FILING DATE: 2003-08-26
 ; PRIOR APPLICATION NUMBER: US 60/401,433
 ; PRIOR FILING DATE: 2002-08-05
 ; PRIOR APPLICATION NUMBER: US 60/449,950
 ; PRIOR FILING DATE: 2003-02-24
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 3418
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-24394-1
 Alignment Scores:
 Pred. No.: 0.00767 Length: 3418
 Score: 149.00 Matches: 213
 Percent Similarity: 34.18% Conservative: 189
 Best Local Similarity: 18.11% Mismatches: 396
 Query Match: 2.63% Indels: 378
 DB: 1 Gaps: 59
 US-09-147-052-3 (1-3261) x PCT-US03-24394-1 (1-3418)
 QY 253 GAGATATACAGTCCATATAA-----TTTAAAGTGACCTTTATTATATAAATATCATTT 306
 Db 1625 GluAsnLeuLysThrSerLysPheLeuLysValLysValHisGluAsnValGlu 1644
 QY 307 CAGACGACCATGAGCGGGGACGATATACAGATCATAATCGCATATACAGATAGG 366
 Db 1645 LysGluThrAlaLysSerProLalaThrCys-----TyrThrAsnGln 1658
 QY 367 AGCGCCGTTTCCATTGAGAGATACGGGATCTAATCGACGGCAAGGAGATGCTCATCT 426
 Db 1659 SerProTyrSerValIleGluAsnSerAlaLeuAlaPheTyrThrSer---CysSerArg 1677
 QY 427 AAAGCA-----AGATACCTTAGAACAATGTA 453
 Db 1678 LysThrSerValSerGlnThrSerLeuLeuGluAlaLysLysTriPleuArgGluGlyIle 1697
 QY 454 TATGTTGAGCGTTTCACAGGATCGGGGAGAAAACAAGTACTCTTAAACCATCAAA 513
 Db 1698 -----PheAspGly-----GlnProGluArg 1704
 QY 514 TTCAACACGCCGGAATCTAGGCGATGGCACAGCACTAATGAGACGTATACCGTGTGGGA 573
 Db 1705 IleAsnThrAlaAsp-----TyrVal 1711
 QY 574 TCACCATGGATATATCGAACGGGAACCTCCGTCAATTGTATAGAGGAAATGGATGCC 633
 Db 1712 GlyAsnTyrLeuTyrGluAsnAsnSer-----AsnSerThrIleAlaGluAsnAspLys 1729
 QY 634 CCTCTGTGTTCCGTATTTCATATTTTGAATGGCCAAATGGCGACATCCGCAACATATCT 693
 Db 1730 AsnHisLeuSerGluLysGlnAspThrTyrLeuSerAsnSerSerMetSerAsnSerTyr 1749
 QY 694 CCATTTTATGGTCTATCCCCACACGAGGCTGCCGAGACCCATGGGATATCCCGCAGGAT 753

Db 1750 SerTyrHisSer-----AspGluValTyrAsnAspSerGlyTyrLeuSerLys 1765
 QY 754 AATTTTCAAACTAGATAGTATTTTCAATGATTG-----GACAAGCGT 801
 Db 1766 Asn-----LysLeuAspSerGlyIleGluProValLeuLysAsnValGluAspGlnLys 1783
 QY 802 CGAAAGCAAGCCTTCCAGTCAAGCGTAACTTTCTCATCATCATCATCATCATCATCATTTGGG 861
 Db 1784 AsnThrSerPheSerLysValIleSerAsnValLysAspAlaAsnAlaTyr----- 1800
 QY 862 TGGGACTGGCTCCAAAATACTACTCTG-----GTATGTTCAATGACTAAGTGAAGAG 915
 Db 1801 -----ProGlnThrValAsnGluAspLysCys----- 1809
 QY 916 GTCGACTGAAATGTTGCGTCAACAGTTAATGGGAGATACAGATTTATGGCCCGTGAACCTT 975
 Db 1810 ValGluGluLeuValThrSerSerProCysLysAsnLysAsnAlaLysLeu 1829
 QY 976 TCGGCAACGTTTATCATAGTCACTGAGTTTGTATCCAAAT-----CGCATC 1023
 Db 1830 Ser-----IleSerAsnSerAsnAsnPheGluValGlyProProAlaPheArgIle 1846
 QY 1024 ATATTAGGACAA-----TGTATTAAACGGGAGGAGAGACCATCGACGAGATATTT 1077
 Db 1847 AlaSerGlyLysIleValCysValSerHisGluThrIleLysLysValLysAspIlePhe 1866
 QY 1078 AGGACAAAATAT-----AATGACAGTCACTCAAGTTGGACAT 1116
 Db 1867 ThrAspSerPheSerLysValIleLysGluAsnAsnGluAsnLysSerLysIleCysGln 1886
 QY 1117 GTACAATATTTCTGCTCTCGGGGATTTATTGTAGCATATCAGCTGTCTTATCCAAA 1176
 Db 1887 ThrLysIleMetAlaGlyCysTyrGluAlaLeuAspAspSerGluLysValSerLysIleSer 1906
 QY 1177 TCCTCTG-----GCTCATATGTACCTCAGAGATTTGTAGAGACACACAGACC 1224
 Db 1907 SerLeuAspAsnAspGluCysSerThrHisSerHisLysValPheAlaAspIleGlnSer 1926
 QY 1225 GATGAGATCTCGACCTGCTGAACATAAGCATGCAATTTATATAAGAAATAT----- 1275
 Db 1927 GluGluIleLeuGlnHisAsnGlnAsnMetSerGlyLeuGluLysValSerLysIleSer 1946
 QY 1276 -----GCTACCTCATTTGTACGATTCGCG 1299
 Db 1947 ProCysAspValSerLeuGluThrSerAspIleCysLysCysSerIleGlyLysLeuHis 1966
 QY 1300 CGAGATATTCGAAATGCACCAATAAGATAAATAACATTAGACGACACACAGCTATATAA 1359
 Db 1967 LysSerValSerSerAla-----AsnThrCysGlyIlePhe 1978
 QY 1360 TCGCATCTCTCT-----GTTCAATTCGCCATGCTCCAA-----TTT 1395
 Db 1979 SerThrAlaSerGlyLysSerValGlnValSerAspAlaSerLeuGlnAsnAlaArgGln 1998
 QY 1396 CTTTATGATCATATACAAACCCATATATATGATATGTTTAGTAGGATTCACAGCTGG 1455
 Db 1999 ValPheSerGluIleGluAspSerThrLysGlnValPheSerLysValLeuPheLysSer 2018
 QY 1456 TCGGAA-----TTGCAAGTAAGAGAACCTT 1479
 Db 2019 AsnGluHisSerAspGlnLeuThrArgGluGluAsnThrAlaIleArgThrProGluHis 2038
 QY 1480 GTTTTATGTCAGCAAGGGATAAAG-----ATTAATCCVAGCCCTACAGCG----- 1524
 Db 2039 LeuIleSerGlnLysGlyPheSerTyrAsnValValAsnSerSerAlaPheSerGlyPhe 2058
 QY 1525 ACTGCAACATTTAGGAGGAGAGTGGCT-----GCCAAG 1557
 Db 2059 SerThrAlaSerGlyLysGlnValSerIleLeuGluSerSerLeuHisLysValLysGly 2078
 QY 1558 ATGTGGGGGATGTCGCTCTGTATCGAGCTGCATATAGATCGGGAATCCGCTCACT 1617

Db 2079 ValLeuGluLupPheAspLeuIleArgThrGluHisSerLeuHisTyrSerProThrSer 2098
QY 1618 TTGCAAAATTCATGCGAGTATACACATCCACTAAT----- 1653
Db 2099 ArgGlnAsnValSerLysIleLeuProArgValAspLysArgAsnProGluHisCysVal 2118
QY 1654 -----ACATGTTATAGCCGACCATTTGGTCTATTTTCATATGAGAA 1695
Db 2119 AsnSerGluMetGluLysThrCys---SerLysGluPheLysLeu-----Ser 2133
QY 1696 AACCAAGGAACATACAGGCAACACTCGGTGAAACAAC-----GAGTTGCTTCCA 1746
Db 2134 AsnAsnLeuAsnValGluGlySerSerGluAsnAsnHisSerIleLysValSerPro 2153
QY 1747 ACCTAGAGGCTGAGAGCCATGCTCGGCTAATCATCTGATATATTTCTGTTGGATCC 1806
Db 2154 TyrLeuSerGlnPheGln-----GlnAspLysGlnGlnLeuValLeuGlyThr 2169
QY 1807 GGTATGCTTTATTTGAAACTATAATTTTGTAAAGATGGTAGACGCTGCC----- 1857
Db 2170 LysValSerLeuValGluAsnIleHisValLeuGlyLysGluGlnAlaSerProLysAsn 2189
QY 1858 -----GATATA 1863
Db 2190 ValLysMetGluIleGlyLysThrGluThrPheSerAspValProValLysThrAsnIle 2209
QY 1864 CAGATTGCTAGCAGATT-----GTCGAG 1887
Db 2210 GluValCysSerThrTyrSerLysAspSerGluAsnTyrPheGluThrGluAlaValGlu 2229
QY 1888 CTTAATCTAACCTGCTAGAGATCGGAAATTTTGCCTTATCCGTTTACACAAAGAA 1947
Db 2230 IleAlaLysAlaPheMetGluAsp-----Asp 2238
QY 1948 GAGTTCGCTGATGTTGTTGATTTATGAGAAGTAGTCGCCGCAATCAACTACAT 2007
Db 2239 GluLeuThrAspSerLysLeuProSerHisAla-----ThrHis 2251
QY 2008 GAACCTTAAATTTTATGACATAACAAGTAATAGAAGTGATGATACAAATATACGCGGGCTG 2067
Db 2252 SerLeu----- 2253
QY 2068 CAGGAATTCGGCTGATGCTATTACTAAAAAGATGCAACCCAAATATGCCAACC 2127
Db 2254 -----PheThrCys-----ProGluAsnGluGluMet 2262
QY 2128 CAATTAGAAGCAGCGCAATG-----GAGTTACAGATCTAATCAATGCT 2172
Db 2263 ValLeuSerAsnSerArgIleGlyLysArgArgGlyGluProLeuIleLeuValGlyGlu 2282
QY 2173 AAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT-----GAAGCT 2220
Db 2283 ProSerIleLysArgAsnLeuLeuAsnGluPheAspArgIleIleGluAsnGlnGluLys 2302
QY 2221 AGTTTATCATCTCTTATAGTGAGCTGAA---ACAGTTAACATACCTTATGCAACA 2277
Db 2303 SerLeuLysAlaSerLysSerThrProAspGlyThrIleLysAsp----- 2317
QY 2278 TTAGAACAACATAAATGGCTAAACATAATTTAGAATCA---GCCATCAACCAAGCTAAT 2334
Db 2318 ---ArgArgLeuPheMetHisHisValSerLeuGluProIleThrCysValProPheArg 2336
QY 2335 ACGGATAAACCAGCTTTTGATATGAAACACCCCAAT----- 2370
Db 2337 ThrThrLysGluArgGlnGluIleGlnAsnProAsnPheThrAlaProGlyGlnGluPhe 2356
QY 2371 TTAGTTGAGCATACAAAGCACTAAACACACTTTAGACAGCTGCTACTAACCCTGAA 2430
Db 2357 LeuSerLysSerHisLeuTyrGluHisLeuThrLeuGluLysSerSerAsnLeu--- 2375
QY 2431 GGTTCGTCATCAACTGCTTATATCAAT-----CGCAATAATTTAGTGGATCTA 2481
Db 2376 AlaValSerGlyHisProPheTyrGlnValSerAlaThrArgAsn----- 2390

QY 2482 TACAATAAAGCTAGTAGTTTAACTATAAATAA-----ACACTAGATACCA 2523
Db 2391 ---GluLysMetArgHisLeuIleThrThrGlyArgProThrLysValPheValProPro 2409
QY 2524 CTAATATGGGGAAACGCTTTTAGATTCTTAATGACATTACTACACTTAATAGATATTAAT 2583
Db 2410 PheLysThrLysSerHisPheHisArgValGluGlnCysValArgAsnIleAsnLeuGlu 2429
QY 2584 AATACGTTTATCAACTATTATGAACAAAAGACTAATGCTGATGATCATTAATAGATTTT 2643
Db 2430 Glu-----AsnArgGlnLysGlnAsnIleAspGlyHisGlySerAspAsp 2444
QY 2644 ATTAANAAGTGATTCAAAATAATGAA-----CAA 2673
Db 2445 SerLysAsnLysIleAsnAspAsnGluIleHisGlnPheAsnLysAsnAsnSerAsnGln 2464
QY 2674 AGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAC---TACAGTTTGT 2730
Db 2465 AlaAlaAlaValThrPheThrLysCysGluGluProLeuAspLeuIleThrSerLeu 2484
QY 2731 GCTTTTAGTCTGCTGATTAACACCCGCTCAATTAT---AAATATGCAAGACGCTTTGG 2787
Db 2485 GlnAsnAlaArgAspIleGlnAspMetArgIleLysLysLysGlnArgGlnArgValPhe 2504
QY 2788 AATGGTGATGAACCTTCAAGTAGAATTTGTCAAACACAGTAATAGTATCACAGATTTCT 2847
Db 2505 -----ProGlnProGlySerLeuTyrLeuAlaLysThrSerThrLeuProArgIleSer 2522
QY 2848 TGGATTATATAGTTAGCTGGA-----ACAACACGAGTAGTACCAATTAGT 2892
Db 2523 LeuLysAlaAlaValGlyGlnValProSerAlaCysSerHisLysGlnLeuTyrThr 2542
QY 2893 TTTAGCAACTATGCTCCATCACTGTTTATTTATTTATTTCCCTTATAGTTGGTTAAAGCA 2952
Db 2543 -----TyrGlyValSerLysHisCysIle----- 2550
QY 2953 GCTGATGCTAATAACGTTGGATTACAAATACAAATAAATAAATAAATGGAATGTTCAACAAGTT 3012
Db 2551 -----LysIleAsnSerLysAsnAlaGluSerPhe 2560
QY 3013 GAGTTTGGCACTTCAACTAGTCGAAATAATATACACTAATCCACTCCAGCAGTTGAT 3072
Db 2561 GlnPheHisThrGluAsp----- 2566
QY 3073 GAGATTAAAGTTGCTAAAATCGTTTATCATAGTTTAAAGATTGGCCAAAACACAATCGAA 3132
Db 2567 -----TyrPheGlyLysGlu----- 2571
QY 3133 TTAAGTGTCCCAACGGGTGAAGAAATATGAATAAAGTTGCGCCAATGATTGGCAACATT 3192
Db 2572 ---SerLeuThrPheThrGlyLysGly-----IleGlnLeuAlaAspGlyGlyTrp 2586
QY 3193 TATCTTAGCTCAATGAATAATGCTGATGAAGATCCCCGGGTACCGT 3240
Db 2587 LeuIleProSerAsnAspGlyLysAlaGlyLysGluGluPheTyrArg 2602

RESULT 7

PCT-US03-26491-192
; Sequence 192, Application PC/TUS0326491
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS T
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/26491
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192

```
; LENGTH: 1261
; TYPE: prt
; ORGANISM: Homo sapiens
PCT-US03-26491-192

Alignment Scores:
Pred. No.: 0.00907 Length: 1261
Score: 147.00 Matches: 136
Percent Similarity: 34.75% Conservative: 109
Best Local Similarity: 19.29% Mismatches: 254
Query Match: 2.59% Indels: 206
DB: 1 Gaps: 34

US-09-147-052-3 (1-3261) x PCT-US03-26491-192 (1-1261)

QY 919 ACTGAATGTGCGTGCACAGTTAAATGGGAGATACAGA-----TTTATGGCCGGTAA 972
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 ThrLysLysLysArgAlaTrpAlaSerGlyGlnLeuSerThrAspIleThrThrSerGlu 28

QY 973 CTTTCGGCAACGTTTATCAGTAATACGACTGAGTTGATCCAAATCGCATCATATTAGGA 1032
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 MetGlyLeuLysSerLeuSerAsnSerIlePheAspProAspTyrIle-----45

QY 1033 CAATGTATTAAACCGAGCGCAGAGCAGCAATCGAGCAGATATTAGCACAAAATATAAT 1092
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 46 -----LysGluLeuValAsn 50

QY 1093 GACAGTCACGTCAGGTGGACATGTACATATTTCTTGGCTCTCGGGGGATTTATTGTA 1152
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 51 AspIleArg---LysPheSerHisIleLeuLeuTyrLeuLysGluAlaIlePheSerAsp 69

QY 1153 GCATATCAGCTGTCTCATATCCAAATCCCTGGCTCATATGTACCTCAGAGAAATTCATGGA 1212
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 CysPheLysGluValIle-----HisIleArgLeuGluGluLeuLeuArg 84

QY 1213 GACAACAGGACCGATGAGATGCTCGACTGGTAAACAATAGCATGCAATTTATAAGAAA 1272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 -----ValLeuLysSerIleMetAsnLysHis-----Gln 94

QY 1273 AATGCTACCTCATGTCACGATTCGGCGGAGATATTGGAATTCGAAATGACCAACATAGAAAATA 1332
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AsnLeuAsnSerVal-----AspLeuGlnAsnAlaIleGluMetLeu---108

QY 1333 ACATTAGACACACACAGCTATTAAATTCGACATCGTCTGTTCAATTCGCCATCGCTCAA 1392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109 -----ThrAlaLysValLys-----AlaValAsnPheThrGluValAsn 121

QY 1393 TTTCTTTATGATCATATACAAACCCATTAATGATATGTTTAGTAGATTGCCACAGCT 1452
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 -----GluGluAsnLysAsnAspLeuPheGlnGluVal-----132

QY 1453 TGGTGCGAATTGCAGATAGAGAACTTGTTTATGGCAGCAGGATTAAGATTAAATCCT 1512
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 -----PheSerSer 135

QY 1513 AGCCTACAGCGATGCAACATTAGGAAGAGAGTGGCTGCAAAAGATGTTGGGGATGTC 1572
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 IleGluThrLeuAlaPheThrPheGlyAsnIleLeuThrAsnPheLeuMetGlyAspVal 155

QY 1573 GCTGCTGTATCGAGCTGC---ACGCTATATAGATGCGGAATCCGTCATCTTGCAAAAT---1626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 GlyAsnAspSerPheLeuArgLeuProValSerArgGluThrLysSerPheGluAsnVal 175

QY 1627 TCTATGCGGATTATCATCCATCAATACATGTTATAGCCGACCATTTGGTCTCTATTTC 1686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 176 SerValGluSerValAspSerSer-----184

QY 1687 TATGGAGAAAACCAAGAAACATACAGGGACAACTCGGTGAAACAAACAGAGTTGCTTCCA 1746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 -----GluLysGlyAsnPhe---SerProLeuGluLeuAspAsnValLeuLeuLys 200

QY 1747 ACGCTAGAGCTGTAGAGCCATGCTCGGCTAAATCATCATGATATTTCTGTTTGGATCC 1806
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 201 AsnThrAspSerIleGlu-----LeuAlaLeu 209

QY 1807 GGTATATGCT---TTATTGAAAACATAT-----AATTTGTTAAGATGTTAGACGCTGCC 1857
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 SerTyrAlaLysThrTrpSerLysThrLysAsnIleValSerTrpValGluLys---228

QY 1858 GATATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAAACCCCTGCTAGAGATCGGAA 1917
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 229 -----LysLeuAsnLeuGluLeuGluSerThrArgAsn 239

QY 1918 ATTTTGCCTTTATCCGTTTACACAAAGAAGAGTTCGTGATGTTGGTATTTGGATAT 1977
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 MetValLysLeuAlaGluAlaThrArg-----ThrAsnIleGlyIleGlnIle 256

QY 1978 GCAGAAGTAGTCGCCGCAATCAATACATCAATCAATTAATTTTATGACATAACAAGATA 2037
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 257 MetProLeu-----GlnSerPheThrAsnAlaLeuLeuAsnAspIleGluSer 274

QY 2038 ATAGAAGTGGATACAAATTAACGGGGCTGCAGGAATTCGGCTGTATGCTCTATTACTAAA 2097
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 HisLeuLeuGlnGlnThrIleAlaAlaLeuGln-----AlaAsn 287

QY 2098 AAAGATGCAAAACCCAAATAATATGGC---CAAAACCAATTAGAAGCAGCCGGAATGAGTGA 2154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 288 LysPheValGlnProLeuLeuGlyArgLysAsnGluMetGluLysGlnArgLysGluIle 307

QY 2155 ACAGAT-----TATAGTGAAGCTGAAACAGTTAACTTAACCTTAATCAAT 2169
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 LysGluLeuTyrLysGlnGlnAsnLysMetLeuGluAlaGluAsnAlaLeuLys 327

QY 2170 GCTAAAGCGATGACATTAAGCTTCTACTACAAGACTATGCCAAGATTGAAGCTAGTTATCA 2229
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 AlaLysLeuLysMetGlnArgGlnAspGluTyrGluLysAlaLysSerSerMetPhe 347

QY 2230 TCTGCT-----TATAGTGAAGCTGAAACAGTTAACTTAACCTTAATCAATCA 2277
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 348 ArgAlaGluGluGlnHisLeuSerSerSerGlyGlyLeuAlaLysAsnLeuAsnLysGln 367

QY 2278 TTGAACAACATAAAATGGCTAAACATAATTTAGAAATCAGCCATCAACCAAGCTAATACG 2337
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 LeuGluLysLysArgArgLeuGluGluGluAlaLeuGlnLysValGluAlaAsp---386

QY 2338 GATAAAACGACTTTTGTATATGAACACCCAAATTTAGTTGAAGCATACAAAGCATAAAA 2397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 -----GluLeuTyrLysValCysVal 393

QY 2398 ACCACTTTAGAACACGCTGCTACTAACCCTGAAGGTTTGTCTCATCACTGCTTATAATCAA 2457
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 ThrAsnValGluGluArgArgAsnAspValGluAsnThrLysArgGluIleLeuAlaGln 413

QY 2458 ATTCGCAATAATTTA-----GTGATCTATACATAAAGCTAGTAGTTTAACTAATA 2511
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 LeuArgThrLeuValPheGlnCysAspLeuThrLeuLysAlaValThrValAsnLeuPhe 433

QY 2512 ACATAGATCCACTAAATGGGGAAACGCTTTTAGATCTTAATGAGATTACTACAGCTAAT 2571
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 434 HisMetGlnHisLeuGlnAlaIleSerLeuAlaIleValArgLeuGlnSerLeuCysGlySer 453

QY 2572 AAGAATATTAAATACGTTTATCACTATTATTAATGAACAAAAGACTAATGCTGATGCATTA 2631
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 AlaLysLeuTyrAspProGlyGlnGluTyrSerGlu-----465

QY 2632 TCTAATAGTTTATTAATAAAGTATTCATAAATATGAACAAAGCTTTTGTAGGACATTT 2691
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 -----PheValLysAlaThrAsnSerThrGluGluLysValAspGly-----480

QY 2692 ACAACGCTAAT-----GTTCAACCTTCAAACTACAGCTTTTGTTCGCTTTT 2736
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 ---AsnValAsnLysHisLeuAsnSerSerGlnProSerGlyPheGly-----495

QY 2737 AGTCTGATGTAAACCGCTCAATTAATAATATCAAGAGGACCGCTTTGGAATGGTAT 2796
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 496 -----ProAlaAsn-----SerLeuGluAspValValArgLeuPro 507
```

QY	2797	GAACCTTCACAGTAGAATT-----CTTTGCAAAACACGAATAGTATTCACAGATGTT	2844
	Asp	:	
Db	508	AppSerSerAsnLysIleGluGluAspArgCysSerAsnSerAlaaspIleThrGlyPro	527
		:	
QY	2845	TCCTTGCAATTATAGTTTACGTGGCAACAACACACAACTACCAATTTAGTTTACCAACTAT	2904
		:	
Db	528	SerPheIleArgSer-----TriPThrPheGlyMetPheSerAsp	540
		:	
QY	2905	GGTCCCATCACTGGT	2919
Db	541	SerGluSerThrGly	545

```

RESULT 8
US-10-648-593-192
; Sequence 192, Application US/10648593
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-192

```

Alignment Scores:	
Pred. No.:	0.00907
Score:	147.00
Percent Similarity:	34.75%
Best Local Similarity:	19.59%
Query Match:	2.59%
DB:	6
Length:	1251
Matches:	136
Conservative:	109
Mismatches:	254
Indels:	206
Gaps:	34

US-09-147-052-3 (1-3261) x US-10-648-593-192 (1-1261)

919	ACTGAAATGTCGGTGCACAGTAAATGGGAGATACAGA-----TTATGGCCCCGTGAA	972
Qy	: : :	
Db	9 ThrLysLysLysArgAlaIatrpAlaSerGlyGlnLeuSerThrAspIlethrThrSerGlu	28
Qy	973 CTTTCGGCAACGTTTATCAGTAATACAGCAGTGTTCGATCCCAATCGCATCATATWAGA	1032
Db	: : : : : : : : :	
Qy	29 MetGlyLeuLysSerLeuSerAsnSerIlePheAspProAspIle	45
Qy	1033 CAATGTATTAAACGCCGAGCAGACGCAATCGACGATATTATTAGGACAAAATATAAT	1092
Db	46 -----LysGluLeuValasn	50
Qy	1093 GACAGTCACGTCAAGGTGGACATGTACAAATATTTCTTGCTCTCGGGGATTTATGTA	1152
Db	51 AspIleArg---LysPheSerHisIleLeuLysIleLeuLysGluAlaIlePheSerAsp	69
Qy	1153 GCATATCAGCCTGCTTCATCCAAATCCCTGGCTCATATGTACCTCAGACAGAAATGTAGAGA	1212
Db	70 CysPheLysGluValIle-----HisIleArgLeuGluGluLeuLeuArg	84
Qy	1213 GACAACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCAATTTATAGAAA	1272
Db	85 -----ValLeuLysSerIleMetAsnLysHis-----Gln	94
Qy	1273 AATGCTACCTCATTTGTCACGATTGCGCGGAGATATTCGAAATGCACCAATAGAAAATA	1332
Db	95 AsnLeuAsnSerVal-----AspLeuGlnAsnAlaIaGluMetLeu---	108
Qy	1333 ACATTAGACACACACCACAGCTATTAAATCGCATCGCTGTTCAAATTCGCCATGCCTCCAA	1392

Db	109	----						----	AlaValAsnPheThrGluValAsn	121
Qy	1393	TTTTCTTTATGATCATATACAAACCCCATATTAATATGATATGTTT	----	AlaValAsnPheThrGluValAsn	1452					
Db	122	-----	----	----	----	----	----	----	GluGluAsnLysAsnAspLeuPheGlnGluVal	132
Qy	1453	TGTCGGCAATTCGAGAAATAGAGAACTGTTTATGCGACGAGGATAAAGATTAACTCT	----	----	----	----	----	----	----	1512
Db	133	-----	----	----	----	----	----	----	----	135
Qy	1513	AGCGCTACACCGAGTGCACATTTAGNAGGAGACTGGCTGCAAAAGATGTTGGGGGATGTC	1572							
Db	136	IleGluThrLeuAlaPheThrPheGlyAsnIleLeuThrAsnPheLeuMetGlyAspVal	155							
Qy	1573	GCTGCTGTATCGAGCTGC-----ACTGCTATAGATCGGAATCCGCTCACTTTTGGCAAAAT---	1626							
Db	156	GlyAsnAspSerPheLeuArgLeuProValSerArgGluThrLysSerPheGluAsnVal	175							
Qy	1627	TCATCGGAGTTATCATATCCATCATATACATGTTATATAGCCGACCATTTGGTCTATTTCAT	1686							
Db	176	SerValGluSerValAspSerSerSer-----	184							
Qy	1687	TATGGAGAAACCAACGGAACATACAGGGGCACTCGGTGCAAAACACGAGTTGCTTCCA	1746							
Db	185	-----GluLysGlyAsnPhe-----SerProLeuGluLeuAspAsnValLeuLeuLys	200							
Qy	1747	ACGCTAGAGCTGTAGACCCATCTCGGCTAATCATCTGATAGATTTCCTGTTTGGATCC	1806							
Db	201	AsnThrAspSerIleGlu-----	209							
Qy	1807	GGTTATGCT---TTATTTGAAACTAT-----AATTTTGTTAAGATGGTAGACGCTGCC	1857							
Db	210	SerTyrAlaLysThrTrpSerLysTyrThrLysAsnIleValSerTrpValGluLys---	228							
Qy	1858	GATATACAGATTCGTAGCACATTTGTCGACCTTAATCTAAACCTGCTAGAAAGATCGGAA	1917							
Db	229	-----LysLeuAsnLeuGluLeuGluSerThrArgAsn	239							
Qy	1918	ATTTTGGCTTTATCCGTTTACACAAAGAGAGTTCGGTGCATGCTGGCTGATTTGGATTAT	1977							
Db	240	MetValLysLeuAlaGluAlaThrArg-----ThrAsnIleGlyIleGlnGluPhe	256							
Qy	1978	GCAGAGTAGCTCGCCGCAATCAACTACATCACTAAATTTTATGACATPAAACAAAGTA	2037							
Db	257	MetProLeu-----GlnSerLeuPheThrAsnAlaLeuLeuAsnAspIleGluSerSer	274							
Qy	2038	ATGAAAGTGAATCAAAATTCACGGGGCTGCAGAAATTCGGCTGTATGCTATTACTPAA	2097							
Db	275	HisLeuLeuGlnGlnThrIleAlaLeuGln-----	287							
Qy	2098	AAAGATGCAAACCCAAATATGCG-----CAAACCCAAATTAGAACGCGCGCAATGGAGTTA	2154							
Db	288	LysPheValGlnProLeuLeuGlyArgLysAsnGluMetGluLysGlnArgLysGluIle	307							
Qy	2155	ACAGAT-----TATAGTGAAGCTGAACACAGTTTAACCAATACCTTATGCAACA	2169							
Db	308	LysGluLeuTrpLysGlnGluAsnLysMetLeuGluAlaGluAsnAlaLeuLysLys	327							
Qy	2170	GCTAAAGCGATGACATTTAGCTTCACATCAAGACTATGCCAAGATTGAAGTAGTTTATCA	2229							
Db	328	AlaLysLeuLeuCysMetGlnArgGlnAspGluTyrGluLysAlaLysSerSerMetPhe	347							
Qy	2230	TCCTGCT-----TATAGTGAAGCTGAACACAGTTTAACCAATACCTTATGCAACA	2277							
Db	348	ArgAlaGluGluGluHisLeuSerSerSerSerGlyLeuAlaLysAsnLeuAsnLysGln	367							
Qy	2278	TTAGAACAACATAAAATGGCTAAACACTAATTTAGTAATCAGCCCATCAACCAAGCTAATACG	2337							
Db	368	LeuGluLysLysArgArgLeuGluGluGluAlaLeuGlnLysValGluGluAlaAsp---	386							
Qy	2338	GATAAAACGACTTTTGATATAATGAACCCCAAAATTTAGTTGACGACATCAACGACCTAAA	2397							

Db 313 LeuLys-----LysLeuThrAspGluMetAlaAlaAspValGlnLysLeuIle 328
QY 2167 AATGTAAGCGATGACATTAGCTTCACTACAGACTATGCCAAGATTGAAGCTATTGA 2226
Db 329 AsnAlaGluThrLeuAsnIleAlaIleLeu----- 338
QY 2227 TCATCTGCTTATAGTGAAGCTCAACACAGTAAACATTAACCTTAATGCAACATTAGAACA 2286
Db 339 -----SerGluAlaGluGlyLeuAsp-----GlnLys 347
QY 2287 CTAAATAAGCTTAAACTAATTTAGAACTCAGCATCAACCAAGCTAATACGGTAAACAG 2346
Db 348 MetLysMetGlnGlnGlnGluLeuGlnLysLeuLysGluLeuGluSerPheArgSer 367
QY 2347 ACTTTTGATATGACACACCAAAATTTAGTTGAAGCATACAAGACACTAAAACCACTTTA 2406
Db 368 SerPheGlnGluGluHisGluLysArgMetGlnAlaGluSerAlaLeuLeuSerGlnGly 387
QY 2407 GAACAACGGTCTACTAACTTCAAA-----GGTTGTCA 2439
Db 388 LysGluLeuAlaGlnSerHisGluGluValGlnArgLeuThrIleGluLysMetAla 407
QY 2440 TCAACTGCTTATAATCAAAATTCGCAATATTTAGTGGATCTATAC----- 2484
Db 408 AsnGluLysLeuAsnGluLeuLysGlnThrLysGluAspLeuHisAspThrValCysGlu 427
QY 2485 -----AATAAGCTAGTAGTTTAAATAACTAAA 2511
Db 428 LeuLysArgAspValGluArgLeuThrGluGlnAsnGlnSerSerGluValLeuIleArg 447
QY 2512 ACACTA--GATCCACTAATGGGGAACGCTTTTAGATTCT--AATCAGATTACTACA 2565
Db 448 GluLeuGlyAspGluIleAsn-----ThrLeuLysAspSerLysAsnGluLeuGlnSer 465
QY 2566 GCTAATAAGAAATAATAATCACTATCACTATTAATAACAAAGAACTAATGCTGAT 2625
Db 466 GluIleLysSerLeuLysSerThrIleSerGlnLeuAsnThrGluLys---AsnAlaAla 484
QY 2626 GCATTATCAAT 2637
Db 485 AlaLeuGlnHis 488

RESULT 10

US-09-820-843B-75
; Sequence 75, Application US/09820843B
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
; TITLE OF INVENTION: PROTEINS USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843B
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1033
; TYPE: PRT
; ORGANISM: M. pneumoniae
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: MG328 homolog
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: gi11674046
US-09-820-843B-75

Alignment Scores:

Pred. No.:	0.00953	Length:	1033
Score:	146.50	Matches:	108
Percent Similarity:	33.99%	Conservative:	83
Best Local Similarity:	19.22%	Mismatches:	182
Query Match:	2.59%	Indels:	189
DB:	5	Gaps:	21

US-09-147-052-3 (1-3261) x US-09-820-843B-75 (1-1033)
QY 1039 ATTAACGCGAG-----GCAGAAAGCAGCAATCGACAGATATTTAGGACAAAATAT 1089
Db 518 IleGluArgGluAsnAsnPheLeuLysThrGluIleLysGlnLeuArgGluThrLysLeu 537
QY 1090 AATGACATGTCAGTCAAGCTTGGACATGATCAATATTTCTTGGCTCTCGGGGATTTATT 1149
Db 538 AsnAspGluAsnThrLysTyr----- 544
QY 1150 GTAGCATATCAGCCTCTTCTATCCAAATCCCTGGCTCATATGTACCTCAGAGAATTGATC 1209
Db 545 -----ArgAsnLeuLeuLysGlnTyr-----GluLeuMet 554
QY 1210 AGAGACAACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCAATTTTAAG 1269
Db 555 ArgAlaAspSerAspAlaLysLeuLysGluLeuGluHisGlnHisLeuAlaHisGln 574
QY 1270 AAAATGCTACCTCATTTGTCACGATTCGCGGAGATATTGGAATGCACCAATAGAAAA 1329
Db 575 HisHisGlnGluGlnLeuAlaGlnLeuGlnArg-----HisAsnGlu 588
QY 1330 ATAACATTTAGACACACACAGCTATTAATCGACATGCTCTGTTCAATTCGCATGCTC 1389
Db 589 AlaLeuValLysGluLeuAspGlnValLysAlaThr----- 600
QY 1390 CAATTTCTTTATGATCATATACAAACCATATTAATGATATGTTTAGTAGGATTCACCA 1449
Db 600 ----- 600
QY 1450 GCTTGTGCGAATTCGAGAATAGAGAACTGTTTATGCGACGAAGGATTAAGATTAAT 1509
Db 601 -----AsnPheGluLeuGlyLeuAlaAlaGlnGlyPheGlu----- 612
QY 1510 CCTAGCGCTACAGCGAGTCAACATTTAGGAGGAGAGTGGCTCAAGATGTTGGGGAT 1569
Db 613 -----GlnGlnLysValValLeuGluGlnLysAsnSerSerLeuLeuAlaSerLeu 639
QY 1570 GTCGCTGCTGTATCGAGCTGCATGCTATAGATGCGGAATCCGCTCACTTTGCAAAATCT 1629
Db 630 GlnAlaAlaGluGluAsnValGlnAlaLeuGlyIleThrAsnSerGluLeuGlnAsnGln 649
QY 1630 ATGCGAGTTATCATCTCCACTAATACATGTTATAGCCGACCATGTTCTATTTCATAT 1689
Db 650 LeuAsnValLeuGluPheThr----- 656
QY 1690 GGAGAAAACCAAGAAACATACAGGACAACTCGGTGAACAAACAGTTGCTTCCAACG 1749
Db 656 ----- 656
QY 1750 CTAGAGCTGTAGAGCCATGCTCGGCTAATCATCGTAGATATTTCTGTTTGGATCCGGT 1809
Db 657 -----HisLysGluLysThrAlaPheAspSer--- 665
QY 1810 TATGCTTTATTGAAACATAATAATTTGTAAGATGTTAGACGCTCCCGCATATACAGATT 1869
Db 666 -----LysThrLeuThrLeuThrLysGlnLeu 675
QY 1870 GCTAGCACATTTGTCAGCTTAATCAACCTCCTAGAGATCGGGAATTTTGGCCTTTA 1929
Db 676 GluGlnThrGlnPheAspLeuSerLeuThr-----GlnGluGlnLeuAlaThr 691
QY 1930 TCCGTTTACACAAAAGAGATTGCGTGTATGTTGGTGTATTGATTATCAGAAGTAGCT 1989
Db 692 PheLysGlnGlnAsnGlnSerLeuThrAsp-----LysLeuMetAla 705
QY 1990 CGCCGCAATCAACTACATGAACCTAAATTTTATGACATAAACAAGTAATAGAGTGGAT 2049
Db 706 SerGluThrGlnLeuAsnHisLeuGlnGlnSerAspGluAsn---LeuThrGlnLeuGln 724
QY 2050 ACAAAATTACGGGGCTGCAGGAATTCGGCTGTATCTCTTACTTAAAAAAGATGCAAC 2109
Db 2109 ----- 2109

Db	725	ThrGlnHisGluLeuLeuGlnGlu	-----SertyrAsnLysLeuGlnAspGlu	740
QY	2110	CCAATAATGCGCAAAACCAATTAGAAGCAGCGGAATGGAGTTAAACAGATCTAATCAAT	2169	
Db	741	AlaAsnHisThrGlnGlnGlnPheHisGlnAlaGlnAsnGluLeu	-----755	
QY	2170	GCTAAAGCGGATGACATTAGCTTCTACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCA	2229	
Db	756	-----AspAlaAlaHisGlnGlnLeuAlaLeuPhe	765	
QY	2230	TCTGCTTATAGTGAAGCTGAACACAGTTAAC	2286	
Db	766	LysGlnAsnAsnGluGluLeuThrAspLysCysSerAsnIleGlnAsnGluLeuHisAsp	785	
QY	2287	CTAAAAATGGCTAAACCTAATTTAGAATCAGCCATCAACCAAGCTAAT	2334	
Db	786	LeuAsnArgValLysThrAsnTrpGluAsnLeuAsnThrGluHisAsnLeuLeuGlnAsp	805	
QY	2335	-----ACGGATAAAGCAGCTTTTGATAATGAACACCCAAATTTAGTT	2376	
Db	806	LysTyrAlaGlnGlnLysGluGlnMetGlnHisGluHisSerAsnLeuAlaGlnIleGln	825	
QY	2377	-----GAGCATACAAAGCAGCTAATAAACACCTTTA	2412	
Db	826	AlaGluHisGluLeuLeuGlnGlnGluSerTyrAsnLysValLysAlaGluLeuAsnGluIle	845	
QY	2413	CGTGCTACTAACCTTGAAGCTTGTCTATCA	2454	
Db	846	GlnIleThrAsnLeuAsnGluAlaAsnAlaGlnTyrGlnAspLeuLeuSerAlaTyrGlu	865	
QY	2455	CAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAACTAAACAA	2514	
Db	866	LeuLeuGlnSerAsn	-----HisAsnLysLeuLysGlnGluLeu---GlnVal	880
QY	2515	CTAGATCCACTAAATGGGGAACGCITTTAGATCTTAATGAGATTACTACAGCTAATAAG	2574	
Db	881	LeuAsnGlnValAsn	-----LeuGluLysGlnGlnLeuAlaGlnLysLeuHis	896
QY	2575	AATATTAAATAATACGTTATCAACTATTAAATGAACAAAAGCTAATGCTGATGCATTATCT	2634	
Db	897	AsnThrHisGlnSerLeuSerGlnThrHisAlaGluLeuThrGlnLeuGlnAlaAlaTyr	916	
QY	2635	AAATAGT	2640	
Db	917	AsnAsn	918	
RESULT 11				
PCT-US03-19153-312				
; Sequence 312, Application PC/TUS0319153				
; GENERAL INFORMATION:				
; APPLICANT: Diversa Corporation				
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND				
; FILE REFERENCE: 09010-250W01				
; CURRENT APPLICATION NUMBER: PCT/US03/19153				
; CURRENT FILING DATE: 2003-06-16				
; PRIOR APPLICATION NUMBER: US 60/389,299				
; PRIOR FILING DATE: 2002-06-14				
; NUMBER OF SEQ ID NOS: 378				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 312				
; LENGTH: 1323				
; TYPE: PRT				
; ORGANISM: Unknown				
; FEATURE:				
; OTHER INFORMATION: Obtained from an environmental sample.				
; FEATURE:				
; NAME/KEY: SIGNAL				
; LOCATION: (1)...(33)				
PCT-US03-19153-312				
Alignment Scores:				
Pred. No.:		0.0126	Length:	1323

[illegible]

Db	811	AsnValThrPheTrpGlyLeuLysAspTyrSerTrpLeuSerThrSerArgSerAsn	830
Qy	2107	AAACCCA-----AATAATGGCCAAACCAAACTTAGAAGCAGCGCGCAATGGAGTTA	2154
Db	831	TrpProLeuLeuPheAspAsnAsnTyrGln	845
Qy	2155	ACGATCATCAATAGCTAAAGCGATGACATAGCTTTCACACAAGACTAT-----	2205
Db	846	TrpAlaIleValGluProSerValLeuProLeuAlaIleAsnLysGlyTyrAlaAsnAsn	865
Qy	2206	-----CCCAAGATTGAAGCTAGTTTATCATCTGCTTAT-----	2238
Db	866	AlaSerAlaArgIleAspGlyValLeuAspArgGluTyrLysGlyAlaIleProIleLys	885
Qy	2239	-----AGTGAAGCTGAACAGTTAAACAATAACCTTAATGCAACATTA	2280
Db	886	IleThrAsnGluSerGlyGlnGluValAlaThrValArgAlaLeuTrpAsnSerSer---	904
Qy	2281	GAACAACATAAATGGCTAAACATAATTTAGATCAGCCATCAACCAAGCTAATACGGAT	2340
Db	905	--GluLeuSerLeuTyrIleSerValAsnAspThrThrIleAspAlaAlaAsnAspLys	923
Qy	2341	AAAACGACTTTT--GATAATGAACCCCAAAATTTAGTTGAAGCA-----TAC	2385
Db	924	ValValValPheValAspGlnAspAsnGlyLysMetProGluLeuLysProAspAspTyr	943
Qy	2386	AAAGCATAAAAACCACTTTAGAACAAAGCTGCTACTAACTTGAAGGTTTGTTCATCACT	2445
Db	944	TrpValSerIleThrArgThrGlyThrLysAlaGlnSerAlaGlnGly-----	959
Qy	2446	GCTTATATCAAAATTCGCATAATTTAGTGGATCTATACATAAAGCTAGTAGTTTAATA	2505
Db	960	-----TyrValLysAspTyrAlaValVal	967
Qy	2506	ACTAAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACA	2565
Db	968	GlnGlnAla-----AsnGly-----TyrValValGlu	976
Qy	2566	GCTAATAAGAAATATAATACGTTATCAACTATTAAATGAACAAAAGACTAATGCTGAT	2625
Db	977	LeuLysLeuLeuIleAsnAsnThrLeu--ThrValAsnSer-----	989
Qy	2626	GCATTATCTAATAGCTTTTATAAAAAGCTGATTCAAAATTAATGAACAAAGTTTGTAGGG	2685
Db	990	-----SerIleGlyPheAspIleAlaIlePheAspAsnGlyValGlnTyrSerTrpAsn	1007
Qy	2686	ACTTTTACAACGCATAAT--GTTCAACCTTCAAACTACAGTTTGTGCTTTTAGTGCT	2742
Db	1008	AspLysThrAsnSerGlnPheIleGluThrAspAsnTyrGlyIleLeuThrMetAlaAsp	1027
Qy	2743	GATGTA-----ACACCCGTC-----	2757
Db	1028	SerValLysPheAlaSerAlaProLysGlyThrAlaIleIleAspAlaGluLeuAspAsp	1047
Qy	2758	AATTAATAAATATGCAACAAGG-----ACCGTT	2784
Db	1048	ThrTrpLysAsnAlaGlnGluIleThrAspThrLysValThrValThrGlyThrVal	1067
Qy	2785	TGGAATGGTGATGAACCTCTCAAGTAGAATCTTGCAACACGAATAGTATACACAGATT	2844
Db	1068	TyrAspSerAlaTyrAlaLysAlaLysMetTrpAspGluAsnSerIleTyrValTyr	1087
Qy	2845	TCTTGGATTATAGTTTA-----GCTGGAACAACACGAAGTACCAATTAGT	2892
Db	1088	AlaIleValTyrAspLeuLeuLeuAsnLysAlaAsnThrAsnProTrpGluGlnAspSer	1107
Qy	2893	TTTAGCAACTATGGTCCATCACTGGTTATTATATTCCTTATAGTTGGTTAAAGCA	2952
Db	1108	IleGluIlePheValAspGluAsnAsnHis-----LysThr	1119
Qy	2953	GCTGATGCTAATAACGTTGGATTACATAACAATAATATGGAATGTTCAACAAGTT	3012
Db	1120	ProTyrTyrGluAsnAspAspValGlnTyrArgValAsnTyrGluAsnThrGln-----	1137

```
QY 3013 GAGTTTGCCTCACTCACTAGTGCATAATAATACAGTAACTCCAACTCCAGCAGTTGAT 3072
    ||| |||: |||: |||
Db 1138 ThrPheGlyThrAsnGlyAlaProGlnAsnPhe-----1148
QY 3073 GAGATTAAAGTTGCTAAATCGTTTATCATCGGT 3105
    ||| |||: |||: |||: |||
Db 1149 ---IleThrAlaThrLysIlelleProAsnGly 1158

RESULT 12
PCT-US03-20322-229
; Sequence 229, Application PC/TUS0320322
; GENERAL INFORMATION:
; APPLICANT: Dana-Farber Cancer Institute, Inc., et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; FILE REFERENCE: DFN-043PC
; CURRENT APPLICATION NUMBER: PCT/US03/20322
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 229
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Bacillus cereus
PCT-US03-20322-229

Alignment Scores:
Pred. No.: 0.0113 Length: 386
Score: 144.50 Matches: 86
Percent Similarity: 38.00% Conservative: 66
Best Local Similarity: 21.50% Mismatches: 148
Query Match: 2.55% Indels: 100
DB: 1 Gaps: 14

US-09-147-052-3 (1-3261) x PCT-US03-20322-229 (1-386)
QY 1888 CTTAATCTAACCGCTGCTAGAGATCGGGAAATTTCCCTTTATCCGTTTACACAAAGAA 1947
    ||| |||: |||: |||
Db 9 LeuLeuValThrAlaValSerThrSerArgPheIleProValSerAlaTyrAlaLysGlu 28
QY 1948 GAGTTCGGTGATGTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2007
    ||| |||: |||: |||
Db 29 GlyGlnThrGluValLysThrValTyrAlaGlnAsnValIleAlaProAsnThrLeuSer 48
QY 2008 GAACCTTAATTTTATGACATAAACAAGTAATAGAGTGGATACAAATACGGGGGCTG 2067
    ||| |||: |||: |||
Db 49 -----AsnSerIleArgMetLeuGlySerGlnSerProLeuIle 61
QY 2068 CAGGAATTCGGCTGTATGCTATTTACTAAAGATGCAACCCAAAT-----2115
    ||| |||: |||: |||
Db 62 GlnAlaTyrGlyLeuIleIleLeuGlnGlnProAspIleLysValAsnAlaMetSerSer 81
QY 2116 -----AATGGCCAAACCAATAGAGCAGCGCGAATGAGTACACATCAATCAAT 2169
    ||| |||: |||: |||
Db 82 LeuThrAsnHisGlnLysPheAlaLysAlaAsnValArgLysIleAspGluTyrAsn 101
QY 2170 GCTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2229
    ||| |||: |||: |||
Db 102 ProLysLeuIleAspLeuAsn-----GlnGluMetMetArgTyrSerThrArgPheAsn 119
QY 2230 TCTGCTTATAGTGAAGCTGAACAGATTAACCTTTAATGCAACATTAGAA-----2283
    ||| |||: |||: |||
Db 120 SerTyrTyrSerLysLeuTyrGluLeuAlaGlyAsnValAsnGluAspGlnGlnAlaLys 139
QY 2284 -----CAACTAAATATGCTAAACATAAT 2307
    ||| |||: |||: |||
Db 140 AlaAspPheMetSerAlaTyrGlyLysLeuGlnValGlnSerIleGlnGluSer 159
QY 2308 TTAGAATCAGCCATCAACCAAGCTAATACGGATAAACGACTTTTGTATATGAACCCCA 2367
    ||| |||: |||: |||
```

```
Db 160 MetGluGlnAspLeuLeuGluLeuAsnArgPheLysThrValLeuAspLysAspSerAsn 179
QY 2368 AATTTA---GTTGAAGCATACAAAGCACTAAACAC-----2400
    ||| |||: |||: |||
Db 180 AsnLeuSerIleLysAlaAspGluAlaIleLysThrLeuGlnGlySerSerGlyAspIle 199
QY 2401 ---ACTTTAGACACACGCTGCTACTAATCACTGAAGTTTGTGCATCACTGCTTATATCAATCA 2457
    ||| |||: |||: |||
Db 200 ValLysLeuArgGluAspIleLysArgIleGlnGly-----Glu 212
QY 2458 ATTGCGCAATAATTTACTGGATCTATACATAAGCTAGTGTAAATA-----2505
    ||| |||: |||: |||
Db 213 IleGlnAlaGluLeuThrThrIleLeuAsnArgProGlnGlnIleLysGlySerIle 232
QY 2506 -----ACTAAACACATAGAT 2520
    ||| |||: |||: |||
Db 233 AsnIleGlyLysGlnValPheThrIleThrAsnGlnThrAlaGlnThrLysThrIleAsp 252
QY 2521 CCCTAAATGGGGGACGCTTTTGTAGATTTCTAATGAGATTACTACAGCTAATAGATATT 2580
    ||| |||: |||: |||
Db 253 PheValSerIleGlyThrLeu-----SerAsnGluIleValAsnAlaAlaAspSerGln 270
QY 2581 AATAATACGTTATCAACTATTATGAACAAAG-----2613
    ||| |||: |||: |||
Db 271 ThrArgGluAlaAlaLeuArgIleGlnGlnLysGlnLeuLeuProLeuIleGln 290
QY 2614 -----ACTAATGCTGATGCTATCTAATAGTTTATTAATAAAGATGATTCAA 2661
    ||| |||: |||: |||
Db 291 LysLeuSerGlnThrGluAlaGluAlaThrGlnIleThrPhe-----ValGlu 306
QY 2662 AATAATGACAAAGTTTGTAGGACTTTTACAAACGCTAATGTTCAACTTCAAACTAC 2721
    ||| |||: |||: |||
Db 307 AspGlnValAsnSerPheThr-GluLeuIleAspArgGlnIleThrThrLeuGluThrLe 326
QY 2722 AGTTTGTGCTTTTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2769
    ||| |||: |||: |||
Db 326 uLeuThrAspTrpLysValLeu-----AsnAsnAsnMetIleGlnIleGlnLysAsnVa 344
QY 2770 GCAAGAAGACCGCTTTGGAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2829
    ||| |||: |||: |||
Db 344 lGluGluGlyThrTyrThrAspSerSerLeuLeuGlnLysHisPheAsnGlnIleLysLys 364
QY 2830 AGTATCACAGATGTTTCTTGGATTATAGTTTACGTGGACCAACACACAGAGTACCAAT 2887
    ||| |||: |||: |||
Db 364 sValSerAsp-----GluMetAsnLysGlnThrAsn 374

RESULT 13
US-10-415-657-10
; Sequence 10, Application US/10415657
; GENERAL INFORMATION:
; APPLICANT: Genix Bioscience GmbH
; TITLE OF INVENTION: Eukaryotic cell division genes and their use in diagnosis and
; FILE REFERENCE: of proliferative diseases
; FILE REFERENCE: CE61773US
; CURRENT APPLICATION NUMBER: US/10/415,657
; CURRENT FILING DATE: 2003-09-08
; PRIOR APPLICATION NUMBER: US 60/246,750
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: C. elegans
US-10-415-657-10

Alignment Scores:
Pred. No.: 0.0154 Length: 1010
Score: 143.50 Matches: 159
Percent Similarity: 35.75% Conservative: 137
Best Local Similarity: 19.20% Mismatches: 273
Query Match: 2.53% Indels: 259
DB: 6 Gaps: 41
```

US-09-147-052-3 (1-3261) x US-10-415-657-10 (1-1010)

1168 CTATCCAAATCCCTGGCTCATATGTACCTCAGAGAATGTATGAGAGACACAGGACCGAT 1227
105 LeuAspLysThrMetAspMetSerLeuGluThrThrIleAsnGluAsnGluThrAla 124
1228 GAGATGCTCGACCTGGTAAACATTAAGCATGCAATTTATTAAGAAAATGCTACTCATTC 1287
125 ArgLeuPheGluThrThrArgAspProThrLeuLeuTyrgLysIleValGluThr 144
1288 TCACGATTTGGGGGAGATATTCGAAATGACCAATAGAAAATAACATTAGACGACACC 1347
145 ThrLysValThrGluArgIleValSerMetPro 159
1348 ACAGCTATTAAATCGCATCG- 1386
160 LeuAlaMetPheAsnThrThrAsnGlnGluAspLysAspMetSerValAspArgSerVal 179
1387 CTCCAATTTCTTTATGATCATATACAAACCCATTAATGATATGTTAGTAGATTGCC 1446
180 Leu---PheThrIleProLysValProLysHis 189
1447 ACAGCTTGGGCGCAATTGCAGAAATAGAGAACTTGTATTATGCGCAGGAGATAAAGATT 1506
189 189
1507 AATCCTAGCGCTACAGCGAGTGAACATTAAGGAGAGAGTGGCTGCAAAAGATGTTGGG 1566
190 190
1567 GATGTCGCTGCTATGCGAGCTGCAGCTGCTATAGATGCGGAATCCGTCACCTTTGCAAAAT 1626
204 SerLysAlaAlaGlyGlnCys 213
1627 TCATGCGAGTTATCATCCATCATACATATACATGTTATAGCGACCATGGTCTATTTCAT 1686
214 ThrMetAsnValPheAsnPheThrAsnLeu 223
1687 TATGGGAAACCAAGGAAACATACAGGAGCAACTCGGTGAAACACAGAGTTGCTTCCA 1746
224 224
1747 ACCTAGAGCGCTGTAGAG---CCATGCTCGCTAAT 1779
238 ThrMetAsnAlaIleArgIleProIleAsnSerAsnValMetProValAspMetAspIle 257
1780 1780
258 ThrGluHisThrLeuLeuGluGluLysLysAsnAspThrPheGlyProSerGlnLeu 277
1816 TTATTTGAAACATATATTTGTTAGATGGTAGAGCTGCGCAT 1860
278 MetAspIleSerAlaProGlnValGlnValAsnAspThrLeuAlaIlePheAsnSerPro 297
1861 1861
298 ArgAspIleCysAsnLysGlyLeuGlyValProGlnAsnLeuIleAsnIleAlaSerAsn 317
1879 TTTGTGCGAGCTTAATCTAACCCCTGCTAGAGATCGGGAATTTGCTTTA 1929
318 ValValProValAspMetAsp---IleThrAspGlnAlaValLeuAsnAlaGluLysLys 336
1930 1930
337 AsnAspGlnPheGluThrSerGlnLeuMetAspIleSerIleProLysValLeuValAsn 356
1978 GCAGAGTACTCGCGCAATCAACATACATGAACTT 2031
357 AspThrMetAlaMetPheAsnSerProLysHisValSerLysSerSerMetAspLeuGlu 376
2032 AAAGTAAATAGAGTG 2073
377 LysThrIleGluAlaAlaAspLysSerThrLysTyProSerIleAlaAspGluValGlu 396
2074 2074
397 AspLeuAspMetAspMetAspIleThrGluGlnGlnProCysGluAlaGlyAsnGlnGln 416
2116 AATGGCCAAACCAATTAAGACGCG---CGATGGAGTTACAGATCTAATCAATGCT 2172
417 AsnAspGlyLeuGlnLeuGlnLysGluAspLeuMetAspIleSerValIleArgAspSer 436
2173 AAAGCGCATG---ACATTAGCTTCACTACAGCATGATCCAGAGTTGAAGTAGTTTA 2226
437 ProAlaValAsnAspThrMetAlaValPheGlnSerProAlaArgValLysIleGlyAla 456
2227 TCATCGCTTATAGTGAAGCTGAA 2250
457 AsnAsnSerIleIleAspSerGlnLysSerIleValPheGlyAspGluMetSerIleAsp 476
2251 2251
477 GluThrGlnAsnAspGlyThrLeuThrLeuProLysSerAsnValGluValThrThr 496
2260 AATAACCTTAATGCAACATTAGAACAACTAAATGCTAAACTAATTTAGATCAAGCC 2319
497 AsnAspValThrSerLeuGluArgGlnGluGluAsnAlaSerGluAsnValSerMet 516
2320 ATCAACCAAGCTAATACG---GATAAAACGACTTTTGTATTAATGAACACCCA 2367
517 IleAsnGluSerSerValHisSerGluIleAspLysLysSerPhe 531
2368 AATTAGCTGAAGCATCAACAGCA---CTAAACCACTTTTAGAACACAGTGTACTAAC 2424
532 MetLeuIleGluGluArgAlaPheMetHisSerMetIleAspValAlaGlnLys 551
2425 CTGTAA---GTTTGTTCATCAACTGCTTATAATCAATTCGAATTCGCAATTTAGTGAT 2478
552 LeuGluAspAspGlySerSerLysThrPro 563
2479 CTATACATAAAGCTAGTCTTATTAATAACTAAA 2511
564 LeuAlaSerGlnSerAlaSerLeuAlaThrLysGluProSerAlaLeuHisAsnSerSer 583
2512 2512
584 AlaThrLeuAsnAsnSerMetGluLeuAspAsnAsnThrLeuLeuLysThrMetGlnIle 603
2560 ACTACAGCTAATAAGATATTAATAACGTTATCAACTATTATTAATGAACAA 2610
604 ThrThrCys 618
2611 AAGACTAATGCTGATGCTATCTAATAGTTTATTAATAAAGTATTAATAAATGA 2670
619 GluLeuAsnSerAsnLysGluGlnGluPheGlyAspGluThrLeuGlnLysAsnAsp 638
2671 CAAGAGTTTGTAGG---ACTTTT 2691
639 ThrSerAsnThrGlyAlaAsnPheThrPheGlnGlyHisAsnGluThrSerGlnIleMet 658
2692 ACAACGCTAATGTTCAACCTTCAACTACAGTCTTCTGCTGCT---TTTAGTGTGATGTA 2748
659 AsnAsnValAspSerGluAlaValAsnThrSerLysIleSerThrThrSerAlaPheAsn 678
2749 ACACCCCTCAATTATTAATATGCA---AGAGAGCCGCTTGTGAATGTGTGAACCT 2802
679 LeuSerIleAsnGlnSerIleSerLysArgArgSerLeuLeuAsnSerAlaArgGlu 698
2803 TCAAGTAGATTTCTTGCACACCAAGATAGTATACAGATGTTCTTGGATTATAGTTTA 2862
699 SerProArgValAlaLeuGluAsnSer 712
2863 GCTGGAAACAAAC 2910
713 AsnGlyGlnThrMetGluAlaLeuThrGluThrArg 729

```

QY 2911 TCACTGGTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAAATACCTT 2970
Db 730 GlnThrSerGlnAspSerMetPro-----SerMetSerLeuAsnAspSer 744
QY 2971 GGATTACAATAACAAATTAATGGAATGTTCAACAAGTTGAGTTGGCCACTTCAACT 3030
Db 745 GlyArgAspIleLeuAlaMetAsnThrSerValArg-----SerProHisLeu 760
QY 3031 ACTGCAATATATACAGCTAAATCCAACTCCAGAGTTGATGAGATTAAAGTTGCTAAA 3090
Db 761 AsnSerSerLysThrAlaAlaProGlyThrProSerLeu----- 773
QY 3091 ATCGTTTTATCAGTTTAAGATTGGCCAAACACAAATCGAATTAAGTGTCCAAACGGT 3150
Db 774 -----MetSerGlnAsn-----ValGlnLeuProProProSerPro 785
QY 3151 GAAGCAATATGAATAGTTGCGCAATGATGTGCAAC-----ATTATCTTAGTCTCAAA 3207
Db 786 GlnPheGluMetProAspPheAspProAlaValAlaAsnValValTyrLeuThrSerGlu 805
QY 3208 GAAATAATGCTGATAAGATCCCC 3231
Db 806 AspProSerThrGluGlnHisPro 813

```

RESULT 14

```

; US-09-581-286A-466
; Sequence 466, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROPHEL, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 466
; LENGTH: 1274
; TYPE: prt
; ORGANISM: Porphyromonas gingivalis
; US-09-581-286A-466

```

```

Alignment Scores:          0.022      1274
Pred. No.:                141.50
Score:

```

```

Percent Similarity: 34.42%      Conservative: 159
Best Local Similarity: 18.27%      Mismatches: 315
Query Match: 2.50%      Indels: 331
DB: 5      Gaps: 51
US-09-147-052-3 (1-3261) x US-09-581-286A-466 (1-1274)
QY 680 TCGGAACATATCTCCATTTTATGGTCTATCCCCACGAGGCTCGCG----- 727
Db 159 SerGlnArgPheIleGluValThrLeuTyrProPheArgTyrAsnProValArgGlnGlu 178
QY 728 -----CAGACCCCATGGATATCCCC 748
Db 179 IleLeuPheAlaLysLysIleGluValThrIleThrPheAsnAspAsnProGln-----Pro 196
QY 749 AGGATAATTTCACAACTAGATAGTATTTCATTCATGGATTGGACAGGCTC----- 802
Db 197 ProLeuGlnLysAsnThr-----GlyIlePheAsnLysValAlaSerSerAlaPheIle 214
QY 803 -----GAAAACAGCCCTTCCAGTCAACGCGTAACCTTCTCATCATCACAC 847
Db 215 AsnTyrGluAlaAspGlyLysSerAlaIleGluAsnAspMetValPheSerArgGlyThr 234
QY 848 ACTTCACAGTTGGTGGGACTGGGCTCCAAAAC----- 881
Db 235 ThrThrTyrIleSerGlyAsn-ValAlaSerAsnLeuProGlnAsnCysAspTyrLeuVal 254
QY 882 ----TACTCGTGTATGTTCAAT-----GACTAAGTGGAAAGAGGTGA 919
Db 254 IleTyrAspAspMetPheAsnValAsnGlnGlnProHisAspGluIleLysArg-Leuc 274
QY 920 CTGAATGTTGCGTGCAACAGTTAATGGGAGATACAGATTTATGGCC-----CGTG 970
Db 274 ysGluHis---ArgAlaPheTyrAsnGly---PheAspValAlaAlaValSerIleLys 292
QY 971 AACTTTCGGCAACGTTTATCAGTATACGACTGAGTTGATCCAAATCGCATCATATTAG 1030
Db 292 spValLeuAsnSerPheProSerAsnAlaThrSerTyr----- 304
QY 1031 GACAATGTATTAAACGCGAGCAGCAGCAATCGAGCAGATATTTAGGACAAATATA 1090
Db 305 -----IleAsnGluThrLysLeuLysAsnPheIleArgSerValTyrA 319
QY 1091 ATGACAGTCAGTCAAG-----GTTGACATGTACATATATTTCTTGCTCTCGGG- 1140
Db 319 snGlnSerAsnAlaLysArgThrLeuAspGlyLysLeuGlyTyrValLeuLeuIleGly 339
QY 1141 -----GGATTATT- 1149
Db 339 ysProLeuSerLysTyrLeuAlaAspThrAspAsnThrLysValProThrSerPheIleH 359
QY 1150 -----GTAGCATATCAGCTGTTCCTATCCAAATCCTCGCTCATATGTACC 1195
Db 359 isAsnValSerLeuIleProSerHisProThrPheGlySerIleCysAlaSerAspTyrP 379
QY 1196 TCAGAGAATTGATGAGACACAGCAGGACCGATGAGATGCTCGACCTG----- 1242
Db 379 hepSerCysVal-----SerProLeuAspThrValGlyAspLeuPheIleGlyArgP 397
QY 1243 -----GTAACAATAAGCATGCAATTTATAGAAAAATGCTACCTCATTTGTACAGA--- 1293
Db 397 heSerValThrAsnAlaHisGluLeuHisAsnLeuIleGluLysThrIleAsnLysGluI 417
QY 1294 -----TTCGCGCGAGATATTCCA-----AATG 1315
Db 417 leSerTyrAsnProIleAlaHisLysAsnIleLeuTyrAlaGluGlyLysGlyCysAspA 437
QY 1316 CACCAATAGAAAATACATATGACACACACAGCTATTAAATCCAGCATCGTCTGT- 1374
Db 437 laProIleLeuArgLeuPheLeuLysGluIleAlaSerGlyTyrThrValAsnSerIleL 457
QY 1375 -----CAATTGCGCATGCTCCAAATTTCTTTATGAT----- 1404

```

Db 457 euLysSerAsnGlnValSerAlaIleAspSerIlePheAspCysLeuAsnAsnGlySerH 477
QY 1405 -----CATATACAAACCCATATAATGATATGTTAGTAGGATGCCACAGCTGGT 1456
Db 477 isHisPheTyrPheAsnThrHis-----GlyMetProThrValTyrPG 491
QY 1457 GCGAATTCAGAAATAGAGNACTGTTTATGCGCAGAGGATGAAGATTAATCCTAGCG 1516
Db 491 lyIleGlyGln-----GlyLeuAspValAsn----- 499
QY 1517 CTACAGCGAGTGCACATTAGCAAGGAGAGTGGCTGCTCAAAAGATGTTGGGGGATGCGCG 1576
Db 500 -----ThrLeuThrAlaArg-----LeuAsnAsnThrSerS 510
QY 1577 CHTGATCGAGCTGCTATGATAGCGGAATCCGTCATCTTGGCAAAATCTATCGCGAG 1636
Db 510 erGlnGlyLeuCysThrSerLeuSerCysSerAlaValAlaAspSerThrIleArgS 530
QY 1637 TTATCATCCACTAATACATCTTATAGCGGACCATGTTCTTATTTTCATATGAGAAA 1696
Db 530 erLeuGlyGlu-----ValLeuThrThrTyrAlaProA 541
QY 1697 ACCAAGGAACATACAGGGACACCTCGGTGAACAAACAGAGTTGCTTCCAAAGCTAGAGG 1756
Db 541 snLyGly---PheSerAlaPheLeuGlyGlySerArgAlaThrGlnTyrAlaValTyrL 560
QY 1757 CHTGATAGCGATCGCGGTATCATCGTAGATATTTCTGTTGGATCCGTTATGCTT 1816
Db 560 euGluGlyProCysProSerGluPheTyrGluTyrLeu-----ProTyrSerL 577
QY 1817 TATTTGAAACATATATTTGTT-----AAGATGGTAGACGCTCCGATATACAGATTG 1870
Db 577 eutyHisAsnLeuSerThrValValGlyGluMetLeuLeuSerIleIleAsnThra 597
QY 1871 CTAGC-----ACATTTGTCGAGCTTAATCTAACCTGCTAGAGAT----- 1911
Db 597 snSerValAspThrTyrSerLysPheAsnPheAsnLeuLeuGlyAspProAlaLeuAsnI 617
QY 1912 -----CGGGAATTTTGC 1924
Db 617 leMetAlaHisGlyMetGluValSerAsnCysIleThrLeuProAsnAsnThrIleIleS 637
QY 1925 CTTATCCGTTTACACAAAGAA-----CAGTTGCGTGTATGTTGCTGTTATGG 1972
Db 637 erSerProIleThrIleLysAsnGlyCysLeuLysIleProGluLysGlyValLeuH 657
QY 1973 ATTATGCA-----GAAGTAGCTCGCC 1993
Db 657 isPheThrAsnAsnGlySerIleGlnValMetSerGlyGlyThrLeuGluIleGlyAsnG 677
QY 1994 GCAATCACTACATGAA-----CTTAATTTTATGAC----- 2025
Db 677 InAlaLysIleSerGlyGluThrGlyAlaAsnProThrPheIleThrValTyrGlyAspG 697
QY 2026 -----ATAACAAAGTAATAGAGTGTATACAAATACGCGGGCTGCAGGATTCG 2077
Db 697 lyLeuAlaIleAsnLysGlnValGluIleAsp---AsnIleAspArgLeuAsnLeuPheS 716
QY 2078 GCTGTATGCTATTACTAAAGAGATGCAACCCCAAAATATGCCAAACCCCAATAGAG 2137
Db 716 erThrHisSerValMetProLysPhe-----HisPheAspSerValLysPheAsnS 733
QY 2138 CAGCGCGAATGGAGTTAAGATCTTAATCAATGCTAAAGCGATGACATTAGCTTACTAC 2197
Db 733 erAlaProLeuTyrThrAsnCysIle----- 742
QY 2198 AGACTATGCCAGATGAGCTAGCTATTCATCTCTTATAGTGAAGCTGAACAGTGA 2257
Db 743 -----ValGluIleSerAsnCysGluPheThrAsnArgSerAspIleIleS 758
QY 2258 ACATAACCTTATGCAACATTAAGAAC----- 2286
Db 758 erLysAsnCysAspLeuSerValGluAsnSerMetPheSerSerGlyIleThrValP 778

QY 2287 --CTAAAAATGGCTAAACATAATTTAGATATCAGCCATCAACCAAGCTAAT---ACGATA 2341
Db 778 heLysProMetAlaThrSerSerIleThrGlyLeuSerThrLysAlaLysIleThrAspA 798
QY 2342 AAACGACTTTTGATATGAACACCCCAATTTAGTTGAAGCATCAACAGCATAAAACCA 2401
Db 798 snThrPhePheAlaThrGlyAsn----- 805
QY 2402 CTTTGAACAACAGCTGCTACTTAACCTTGAAGGTTTGTCACTCAACTGCTTATTAATCAATTC 2461
Db 806 --PheAlaTyrHisIleThrAsnThrProGlyLeuThrAlaThrSerAsnAlaIleL 825
QY 2462 GC---AATTAATTTAGTGGATCTATACATATAAGCTAGTAGTTAATACTAAACACTAG 2518
Db 825 ysLeuAspAsnIleProGluTyrTyr----- 833
QY 2519 ATCCACTAATGGGGGACGCTTTTATAGATCTTAATGATTAATGAGATTACTACAGCTAATAAGNATA 2578
Db 834 -----IleSerGlyAsnLysIleValAsnCysAspGluAlaLeuValLeuAsnAsnSerG 852
QY 2579 TTAATAATACGTTTATCAACTATTAATGAACAAAGACT----- 2616
Db 852 lyAsnArgThrAsnArgLeuHisAsnIleThrArgAsnValIleLysAsnCysArgIleG 872
QY 2617 -----AATGCTGATGCTATCTATTAATGATTTTA 2644
Db 872 lySerThrLeuTyrAsnSerTyrGlyIleTyrAsnArgAsnLysIleSerAsnAsnHisI 892
QY 2645 TTAATAAAGCTGATTCAAATATGAACAAGTTTGTAGGACTTTTACAAAGCTTAATG 2704
Db 892 leGlyValArgLeuLeuAsnAsnSerCysPheTyr-----PheAspAsnAlaProV 909
QY 2705 TTCAACCT-----TCAAACTACAGTTTGTGCTTTAGTG 2740
Db 909 alIleAsnGluGluAspLysGlnThrPheIleSerAsnArgThrTyrGlnLeuTyrSerS 929
QY 2741 CTGATGTAACA---CCGTCATATTAATATGCAAGAGCCGTTTGAATGGTATG 2797
Db 929 erAsnGlyThrPheProLeuAsnPheHisTyr-----AsnSerLeuGlnGlyGlyAsp- 946
QY 2798 AACCTTCAAGTAGAATTCITGCAACACAGATAGTATCATCAGATGTTCTTGGCATTTATA 2857
Db 947 -----ThrAsp-----ThrTrpIleTyrA 953
QY 2858 GTTACTGGAACAAACAGCACTACCAATTTAGTTTGTAGCAACTATGTTCCATCACTG 2917
Db 953 snAspThrTyrThrAsnArgTyrIleAspValSerAsnAsnHisTyrGlyAsnAsnAsp- 972
QY 2918 GTTATTTATTTCCCTTATAGTTCGTTAAAGCAGCTGATGCTATATACGTTGGATTAC 2977
Db 973 -----LeupheAspProAsnGlnValPheAsnThrProAspLeu---PheIleTrpIleP 990
QY 2978 AATAC-----AAATTAATAATGAAATGTTTCAACAAGTTGAGT 3016
Db 990 roPheThrAspGlyLeuProAsnGlyArgSerGlyAsnSerSerAlaGluAlaValGluP 1010
QY 3017 TTGCCACTTCA 3027
Db 1010 heGlnThrAla 1013

RESULT 15

US-09-581-286A-465
; Sequence 465, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHLE, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.

TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
 FILE REFERENCE: 4137-3
 CURRENT APPLICATION NUMBER: US/09/581,286A
 CURRENT FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: PCS/AU98/01023
 PRIOR FILING DATE: 1998-12-10
 PRIOR APPLICATION NUMBER: AU PP 0839
 PRIOR FILING DATE: 1997-12-10
 PRIOR APPLICATION NUMBER: AU PP 1182
 PRIOR FILING DATE: 1997-12-31
 PRIOR APPLICATION NUMBER: AU PP 1846
 PRIOR FILING DATE: 1998-01-30
 PRIOR APPLICATION NUMBER: AU PP 2264
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: AU PP 2911
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: AU PP 3128
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: AU PP 3338
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: AU PP 3654
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: AU PP 4917
 PRIOR FILING DATE: 1998-07-29
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 721
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 465
 LENGTH: 1282
 TYPE: PRT
 ORGANISM: Porphyromonas gingivalis
 US-09-581-286A-465

Alignment Scores:
 Pred. No.: 0.022 Length: 1282
 Score: 141.50 Matches: 180
 Percent Similarity: 34.42% Conservative: 159
 Best Local Similarity: 18.27% Mismatches: 315
 Query Match: 2.50% Indels: 331
 DB: 5 Gaps: 51

US-09-147-052-3 (1-3261) x US-09-581-286A-465 (1-1282)

QY	680	TCGCGAACATATCCATTTTATGGTCTATCCCCACGAGGCTGCCG	-----727
DB	167	SerGlnArgPheIleGluValThrLeuTyrProPheArgTyrAsnProValArgGlnGlu	186
QY	728	-----CAGAACCCCATGGGATATCCCC	748
DB	187	IleLeuPheAlaLysLysIleGluValThrIleThrPheAspAsnProGln	-----Pro 204
QY	749	AGGATAATTCACAACTAGATAGCTATTTTCAATGGATTGGACAGCGTC	-----802
DB	205	ProLeuGlnLysAsnThr-----GlyIlePheAsnLysValAlaSerSerAlaPheIle	222
QY	803	-----GAAAGCAGCGCTCCAGTCAAGCTAACTTCTCATCATCACATCAC	847
DB	223	AsnTyrGluAlaAspGlyLysSerAlaIleGluAsnAspMetValPheSerArgGlyThr	242
QY	848	ACTTCACAGTTGGTGGGACTGGGCTCCAAAAC	-----881
DB	243	ThrThrTyrIleSerGlyAsn-ValAlaSerAsnLeuProGlnAsnCysAspTyrLeuVa	262
QY	882	-----TACTCGGTATGTTCAAT	-----
DB	262	IleTyrAspAspMetPheAsnValAsnGlnProHisAspGluIleLysArg-LeuC	282
QY	920	CTGAATGTTGGTGGTCAACAGTTAAATGGAGATACAGATTTATGGCC	-----CGTG 970
DB	282	ysGluHis---ArgAlaPheTyrAsnGly---PheAspValAlaAlaValSerIleLysA	300
QY	971	AACTTTCGGCAACGTTTACAGTAATACGACTGAGTTGATCCAAATCGCATATATAG	1030

```

QY 1912 -----CGGGAATTTTGC 1924
Db 625 leMetalaHisGlyMetGluValSerAsnCysIleThrLeuProAsnAsnThrIleIles 645
QY 1925 CTTTATCCGTTTACACAAAGAA-----GAGTTGGGTGATGTTGGTGTATTGG 1972
Db 645 erSerProIleThrIleLysAsnGlyCysLeuLysIleProGluLysGlyValLeuH 665
QY 1973 ATTATGCA-----GAAGTAGCTCGCC 1993
Db 665 isPheThrAsnAsnGlySerIleGlnValMetSerGlyGlyThrLeuGluIleGlyAsnG 685
QY 1994 GCAATCACTACATGNA-----CTTAATTTTATGAC---- 2025
Db 685 InAlaLysIleSerGlyGluThrGlyAlaAsnProThrPheIleThrValTyGlyAspG 705
QY 2026 -----ATAACAAAGTAATAGAGTGCATACAAATTCACGGGGCTGCAGGAATTCG 2077
Db 705 lYleuAlaIleAsnLysGlnValGluIleAsp---AsnIleAspArgLeuAsnLeuPheS 724
QY 2078 GCTGTATGCTATTACTAAAGAAAGATGCAAAACCCAAATAATGCCAAACCCAAATTAGAAG 2137
Db 724 erThrHisSerValMetProLysPhe-----HisPheAspSerValLysPheAsnS 741
QY 2138 CAGCGCGAATGAGTTACAGATCTAATCAATGCTTAACCGATGACATTAGCTTCACTAC 2197
Db 741 erAlaProLeuTyThrThrAsnCysIle----- 750
QY 2198 AAGACTATGCCAAGATGAAGCTAGTTTATCATCTGCTTATAGTAGAAGCTGAACACAGTTA 2257
Db 751 -----ValGluIleSerAsnCysGluPheThrAsnArgSerAspIleIles 766
QY 2258 ACAATAACCTTAATGCAACATTAGAACAA----- 2286
Db 766 erLysAsnCysAspLeuSerValGluAsnSerMetPheSerSerSerGlyIleThrValP 786
QY 2287 --CTAAAAATGCTAAACATAATTTAGATACCCATCAACCAAGCTAAT---ACGGATA 2341
Db 786 heLysProMetAlaThrSerSerIleThrGlyLeuSerThrLysAlaLysIleThrAspA 806
QY 2342 AAACGACTTTTGATATGAACACCCCAAAATTTAGTTGAAGCATACAAGCACATAAAACCA 2401
Db 806 snThrPheAlaThrGlyAsn----- 813
QY 2402 CTTTAGAACACGTCGTACTAACCTTGAAGTTTGTATCACTGCTTATATCAATCAATTC 2461
Db 814 --PheAlaTyHHisIleThrAsnThrProGlyLeuThrAlaThrSerAsnAlaIleL 833
QY 2462 GC---AATATTTAGTGGATCTATACAATAAAGCTAGTAGTTTATAACTAAACACTAG 2518
Db 833 ysLeuAspAsnIleProGluTyTy----- 841
QY 2519 ATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAGAATA 2578
Db 842 -----IleSerGlyAsnLysIleValAsnCysAspGluAlaLeuValLeuAsnAsnSerG 860
QY 2579 TTAATAATACGTATCACTATTAATCAACAAAGAGACT----- 2616
Db 860 lyAsnArgThrAsnArgLeuHisAsnIleThrArgAsnValIleLysAsnCysArgIleG 880
QY 2617 -----AATGCTGATGATTATCTAATAGTTTAA 2644
Db 880 lySerThrLeuTyAsnSerTyGlyIleTyAsnArgAsnLysIleSerAsnAsnHisI 900
QY 2645 TTAATAAAGTATCAATAATAATGAACAAAGTTTGTAGGACTTTTACAAACGCTAATG 2704
Db 900 leGlyValArgLeuLeuAsnAsnSerCysPheTy-----PheAspAsnAlaIle 917
QY 2705 TTCAACCT-----TCAAACCTACAGTTTGTGCTTTTAGTG 2740
Db 917 alIleAsnGluGluAspLysGlnThrPheIleSerAsnArgThrTrpGlnLeuTySerS 937

```

Search completed: October 8, 2003, 19:18:03
 Job time : 132.456 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:53:53 ; Search time 986.746 Seconds
(without alignments)
6014.205 Million cell updates/sec

Title: US-09-147-052-3

Perfect score: 5665

Sequence: 1 atgcactatttagcgga.....gaccgggtacattttataa 3261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 11457514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09147052/runat_08102003_154343_29901/app.query.fasta_1.5980
-DB=Pending_Patents_AA_Main -QFMT=fastan -SURFIX=rapm -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052.ecgn.1.1.2602.drunat.08102003_154343_29901 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5619	99.2	1086	15	US-09-147-052-4	Sequence 4, Appl
2	3601	63.6	865	3	US-07-722-860-13	Sequence 13, Appl
3	3601	63.6	865	6	US-08-213-449A-13	Sequence 13, Appl
4	3601	63.6	865	6	US-08-213-449B-13	Sequence 13, Appl
5	2015	35.6	456	15	US-09-147-052-2	Sequence 2, Appl
6	2008	35.4	456	24	US-09-901-572A-3	Sequence 3, Appl
7	1929.5	34.1	384	27	US-10-131-591A-12	Sequence 12, Appl
8	1777.5	31.4	868	12	US-08-804-439-21	Sequence 21, Appl
9	1777.5	31.4	868	17	US-09-301-390-21	Sequence 21, Appl
10	1777.5	31.4	868	17	US-09-338-326-21	Sequence 21, Appl
11	1777	31.4	357	27	US-10-131-591A-23	Sequence 23, Appl
12	1749	30.9	933	3	US-07-805-524-2	Sequence 2, Appl
13	1743	30.8	943	4	US-08-096-183D-4	Sequence 4, Appl
14	1743	30.8	943	19	US-09-521-738-14	Sequence 14, Appl
15	1739.5	30.7	891	9	US-08-541-878-6	Sequence 6, Appl
16	1736	30.6	904	1	PCT-US03-11231-18	Sequence 18, Appl
17	1736	30.6	904	27	US-10-121-988-18	Sequence 18, Appl
18	1736	30.6	904	28	US-10-200-562-18	Sequence 18, Appl
19	1736	30.6	904	28	US-10-237-551-18	Sequence 18, Appl
20	1719.5	30.4	885	9	US-08-541-878-8	Sequence 8, Appl
21	1714.5	30.3	885	12	US-08-804-439-23	Sequence 23, Appl
22	1714.5	30.3	885	17	US-09-301-390-23	Sequence 23, Appl
23	1714.5	30.3	885	17	US-09-338-326-23	Sequence 23, Appl
24	1702.5	30.1	903	12	US-08-804-439-22	Sequence 22, Appl
25	1702.5	30.1	903	17	US-09-301-390-22	Sequence 22, Appl
26	1702.5	30.1	903	17	US-09-338-326-22	Sequence 22, Appl
27	1696.5	29.9	904	32	US-60-412-956-12	Sequence 12, Appl
28	1695	29.9	846	5	US-08-123-456-142	Sequence 142, App
29	1695	29.9	846	16	US-09-297-477A-142	Sequence 142, App
30	1695	29.9	846	25	US-09-994-404-142	Sequence 241, App
31	1695	29.9	896	5	US-08-123-456-241	Sequence 241, App
32	1695	29.9	896	16	US-09-297-477A-241	Sequence 241, App
33	1695	29.9	896	25	US-09-994-404-241	Sequence 241, App
34	1664.5	29.4	787	5	US-08-123-456-79	Sequence 79, Appl
35	1664.5	29.4	787	16	US-09-297-477A-79	Sequence 79, Appl
36	1664.5	29.4	787	25	US-09-994-404-79	Sequence 79, Appl
37	778.5	13.7	907	4	US-08-048-978-2	Sequence 2, Appl
38	778.5	13.7	907	12	US-08-804-439-19	Sequence 19, Appl
39	778.5	13.7	907	17	US-09-301-390-19	Sequence 19, Appl
40	778.5	13.7	907	17	US-09-338-326-19	Sequence 19, Appl
41	778.5	13.7	907	28	US-10-223-538-2	Sequence 2, Appl
42	758.5	13.4	830	26	US-10-055-364-45	Sequence 45, Appl
43	756	13.3	906	32	US-60-435-549-11	Sequence 11, Appl
44	754	13.3	865	1	PCT-US00-18647-24	Sequence 24, Appl
45	754	13.3	865	26	US-10-055-364-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-147-052-4

; Sequence 4, Application US/09147052

; GENERAL INFORMATION:

; APPLICANT: SAITOH, Shuji

; APPLICANT: TSUZAKI, Yoshinari

; APPLICANT: YANAGIDA, Noboru

; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,

; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE

; FILE REFERENCE: 981167

```

; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
; US-09-147-052-4

```

Alignment Scores:

```

Pred. No.: 0 Length: 1086
Score: 5619.00 Matches: 1086
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.19% Indels: 0
DB: 15 Gaps: 0

```

US-09-147-052-3 (1-3261) x US-09-147-052-4 (1-1086)

```

QY 1 ATGCACATATTTAGGCGGAATTGCATATTTTCTTATAGTTATCTATATGTTACGCAAC 60
DB 1 MethisTyrPheargarganscysillePheLeulleValilleLeuTyrGlyThrAsn 20
QY 61 TCATCTCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGGAGAGTCTACGTTTATCTTGTCTCCCGCCACAGTGGTTCAACCGTGATCCGCTCA 180
DB 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValilleArgLeu 60
QY 181 GAACCCCGCGAAATGTCCCGAACCTAGAAAAGCCAGGAGTGGGCTGAAGGAATCGCG 240
DB 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlylleAla 80
QY 241 ATATTATTTAAGAGAATATCAGTCCATATAATTTAAAGTGACGCTTTATATAAAAT 300
DB 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCACAGCAGACATGGGCGGAGCAGATATAGACAGATCACTAATCGATATACA 360
DB 101 IleileGlnThrThrTrpThrGlyThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGCAGCCCGTTTCCATTGAAGAGATCAGCGATCTAATCGACGCCAAAGGAAGATGC 420
DB 121 AspArgThrProValSerIleGluGluIleThrAspLeulleaspGlyLysGlyArgCys 140
QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGATATGTTGAAGCGTTTGACAGGGATGG 480
DB 141 SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheaspargaspAla 160
QY 481 GGAGAAAACACAGTACTTCTAAACCATCAAAATTCACACGCCCGCAATCTAGGGCATGG 540
DB 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACACGACTAANTCAGACGTATACCGTGTGGGATCACCATGATATATATGATATATG 600
DB 181 HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCCGTCATTTGATAGTAGAAGAAATGGATGCCGCTCTGTGTTTCCGTATTCAATTTT 660
DB 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCATGGCCCAANTGGGACATCGGAACATATCTCCATTTTATGGTCTATCCCGCCACAGAG 720
DB 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProGlu 240
QY 721 GCTGCCGCGAAGCCCATGGGATATCCCGAGGATATTTTCAACAACATAGATACCTATTT 780

```

```

DB 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelysGlnLeuaspSerTyrPhe 260
QY 781 TCAATGGATTGACAAGCGTCGAAAAGCAAGCCTTCCAGTCAAGGTACTTCTTCATC 840
DB 261 SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeulle 280
QY 841 ACATCACACTTCACAGTTGGTGGGACTGGGCTCCAAAAAACTACTCGTGTATGTTCAATG 900
DB 281 ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet 300
QY 901 ACTAAGTGAAGAGGTGACTGAAATGTTGCGTGCACACAGTTAATGGGAGACATACAGATT 960
DB 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
QY 961 ATGGCCCGTGAACCTTTTCGCAACGTTTATCAGTAATACGACTGAGTTGATCCAAATCGC 1020
DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrLupPheaspProAsnArg 340
QY 1021 ATCATATTTAGGACATGTATTAAACCGGAGGCGAGAGCAATCGAGCAGATATTTAGG 1080
DB 341 IleileLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
QY 1081 ACAAAATATAATCACAGTCACGTCACGTCAGGTTGGACATGTACAATATTTCTTGGCTCTCGG 1140
DB 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
QY 1141 GGATTTATTGTAGCATATCAGCCTGTCTTATCCAAATCCCTGGCTCATATGTACCTCAGA 1200
DB 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
QY 1201 GAATTCATGAGACACAACAGGACCGATGAGATGCTCGACCTGTAAACAATAAGCATGCA 1260
DB 401 GluLeuMetArgaspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATAAGAAAATGCTACCTCATTTGTCAGGATTCGGGAGATATTCGAAATGACCA 1320
DB 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro 440
QY 1321 ATAGAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTCTCAATTC 1380
DB 441 AsnArgLysIleThrLeuaspAspThrThrAlaIleLysSerThrSerSerValGlnPhe 460
QY 1381 GCCATCTCCAATTTCTTTATGATCATATACAACCCCATTAATGATATGTTTAGTAGG 1440
DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
QY 1441 ATTGCCACAGCTTGGTGGAAATTCAGAAATAGAGAACTTGTCTTTATGGCACGAAGGATA 1500
DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle 500
QY 1501 AAGATTAACTTAGCGCTACAGCGAGTGCACATATTAGGAAGAGAGTGGCTGCAAGATG 1560
DB 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
QY 1561 TTGGGGGATGTCCTGCTGTATCGAGCTGACGTGCTATAGATCGGGAATCCGTCACCTTG 1620
DB 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleaspAlaGluSerValThrLeu 540
QY 1621 CAAATTTCTATCGGAGTTATCACATCCACTAATACATGTTATAGCCGACCATTTGGTTCTA 1680
DB 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTTCATATGGAAAACCAAGGAAACATACAGGACAACTCGGTGAAAAACAACAGAGTTG 1740
DB 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTTCAACCGCTAGAGCGCTAGAGCCATGCTCGGCTAAATCATCATCGTATGATTTCTGTTT 1800
DB 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgArgTyrPheLeuPhe 600
QY 1801 GGATCCGGTTATGCTTTTATTGAAAACTATATTTTGTAAAGATGCTAGACGCTGCCGAT 1860

```

Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAsp 620
 QY 1861 ATACAGATTGCTAGACACATTTTCGAGCTTAATCAACCTGCTAGAGATCGGAAAT 1920
 Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuGluAspArgGluIle 640
 QY 1921 TTGGCTTTATCCGTTTACACAAAGAAGAGTTGCGTGATGTTGGTGATTTGATGATATGCA 1980
 Db 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspPyrAla 660
 QY 1981 GAAGTAGCTCGCCGAATCACTACATGAACCTTAATTTTATGACATACAAAGTAATA 2040
 Db 661 GluValAlaArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
 QY 2041 GAAGTGGATACAAATTCGCGGGCTGCAGGAATTCGCGTGATGCTATTACTAAATA 2100
 Db 681 GluValAspThrAsnTyrAlaGlyLeuGlnGluPheGlyCysMetSerIleThrLysLys 700
 QY 2101 GATGCAACCCAAATAATGCGCAACCCCAATTTAGAGCAGCGCAATGAGTTAACAGAT 2160
 Db 701 AspAlaAsnProAsnAsnGlyGlnThrGlnLeuGluAlaAlaArgMetGluLeuThrAsp 720
 QY 2161 CTAATCAATGCTAAAGCGATGACATTAGCTTCACATACAGACTATGCCAAGATTGAAGCT 2220
 Db 721 LeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAla 740
 QY 2221 AGTTTATCATCTGCTTATAGTAGAGCTGAAACAGTTTAAACATTAATCAACATTA 2280
 Db 741 SerLeuSerSerAlaTyrSerGluAlaGluThrValAsnAsnAsnLeuAsnAlaThrLeu 760
 QY 2281 GAACAACTAAATAATGGCTAAACTAATTTAGATCAGCCATCAACCAAGCTAATACGGAT 2340
 Db 761 GluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsnThrAsp 780
 QY 2341 AAACGACCTTTTGATTAAGAACCCCAATTTAGTTGAAGCATACAAAGCCTAAACACC 2400
 Db 781 LysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeuLysThr 800
 QY 2401 ACTTTAGAACACAGCTGCTACTACCTTGAGCTTGTGCATCACTGCTTATTAATCAAT 2460
 Db 801 ThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsnGlnIle 820
 QY 2461 CCAATAATTTAGTGGATCTATACAAATAAGCTAGTGTAACTAACTAAACACTAGAT 2520
 Db 821 ArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThrLeuAsp 840
 QY 2521 CCACATAATGGGGGAACCGCTTTTAGATTCTAATGAGATTACTACAGCTAATTAAGAAAT 2580
 Db 841 ProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLysAsnIle 860
 QY 2581 AATAATAGCTTATCAACTATTAAATGAACAAAGACTAATGCTGATGCTATTCTAATAGT 2640
 Db 861 AsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSerAsnSer 880
 QY 2641 TTTTATTAATAAAGTATCAATAATGAACAAAGCTTTGTAGCGACTTTTCAACACCT 2700
 Db 881 PheIleLysLysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThrAsnAla 900
 QY 2701 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGTGCTGATGCTACACCCGCTCAAT 2760
 Db 901 AsnValGlnProSerAsnTyrSerPheValAlaPheSerAlaAspValThrProValAsn 920
 QY 2761 TATAATATGCAAGAGACCGCTTTGGATGCTGATGAACCTTCAAGTAGAATTCCTGCA 2820
 Db 921 TyrLysTyrAlaArgThrValTrpAsnGlyAspGluProSerSerArgIleLeuAla 940
 QY 2821 AACACGAATGATCACAGATCTTCTGGATTTATAGTTTGTAGCTGGAGCAACACAGAG 2880
 Db 941 AsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsnThrLys 960
 QY 2881 TACCAATTTAGTTTGTAGCAACTATGCTCATCACTGTTTATTTATTTTCCCTTATAG 2940
 Db 961 TyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 980

QY 2941 TTGGTTAAAGCAGCTGATGCTTAATACGTTGATTACAAATACAAATTAATATGAAAT 3000
 Db 981 LeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGlnTyrLysLeuAsnAsnGlyAsn 1000
 QY 3001 GTTCAACAGTTGAGTTTCCCACTCACTAGCTAGTGCAGAAATAACTACTACACTTAATCCAAT 3060
 Db 1001 ValGlnGlnValGluPheAlaThrSerThrSerAlaAsnAsnThrThrAlaAsnProThr 1020
 QY 3061 CCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTTTAAGATTGGCCAA 3120
 Db 1021 ProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPheGlyGln 1040
 QY 3121 ACACAACTCGAATTAAGTCTTCCACGGGTGAAGGAAATATGAATAAACTTGGCCCAATG 3180
 Db 1041 AsnThrIleGluLeuSerValProThrGlyGluGlyAsnMetAsnLysValAlaProMet 1060
 QY 3181 ATTGCAACATTTACTTAGCTCAATGAATGAATAATGCTGATAAGATCCCCGGTACCCT 3240
 Db 1061 IleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLysIleProGlyTyrArg 1080
 QY 3241 CGACCCGGTACATTTT 3258
 Db 1081 ArgProGlyThrPheLeu 1086

RESULT 2

US-07-722-860-13
 ; Sequence 13, Application US/07722860
 ; GENERAL INFORMATION:
 ; APPLICANT: NAZERIAN, Keyvan
 ; APPLICANT: LEE, Lucy F.
 ; APPLICANT: YANAGIDA, Noboru
 ; APPLICANT: OGAWA, Ryohel
 ; APPLICANT: LI, Yi
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
 ; PROTECTION AGAINST MAREK'S DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: 301 North Washington Street
 ; City: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/722,860
 ; FILING DATE: 19910628
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M.
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 1644-101P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 241-1300
 ; TELEFAX: (703) 241-2848.
 ; TELEX: 248345
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 865 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-722-860-13

Alignment Scores:
 Pred. No.: 0 Length: 865
 Score: 3601.00 Matches: 719
 Percent Similarity: 87.81% Conservative: 16

Best Local Similarity: 85.90%		Mismatches: 60	
Query Match: 63.57%		Indels: 42	
DB: 3		Gaps: 8	
US-09-147-052-3 (1-3261) x US-07-722-860-13 (1-865)			
QY	1	ATGCACTATTTTAGCGGAATTCGATATTTTCTTATAGTATTCATATGTTACGAAAC	60
Db	1	MetHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn	20
QY	61	TCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTGTGTTTCGAGCGCTCAGTGTCT	120
Db	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40
QY	121	GAGGAAGAGTCTACGTTTTATCTTTGTCTCCCCACAGTGGTTCACACCGTGATCCGCTCA	180
Db	41	GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu	60
QY	181	GAACCGCGCGAAATGTCGCAACCTAGAAAAGCCACCGAGTGGGTGAAGGAATCGCG	240
Db	61	GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla	80
QY	241	ATATTATTAAAGAGAATATCAGTCCCATATATAAATTTAAAGTGAGCGCTTATTATAAAAT	300
Db	81	IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn	100
QY	301	ATCATTCAGAGCAGCATGCGCGGGAGCAGACATATAGACAGATCACTAATCGATATACA	360
Db	101	IleIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr	120
QY	361	GATAGAGCCCGTTTCCATTGAAGAGATCACGGATCTAATCCAGCGCAAGGAAGATGC	420
Db	121	AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys	140
QY	421	TCATCTAAGCAAGATACCTTAGAAAATGTATATGTTTGAAGCGTTTGACAGGGATGCG	480
Db	141	SerSerLysAlaArgTyrLeuArgAsnValTyrValGluAlaPheAspArgAspAla	160
QY	481	GGAGAAAACAGTACTTCTAAAACCATCAAAATTCACACGCCCGCAATCTAGGGCATGG	540
Db	161	GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180
QY	541	CACACGACTAATGAGAGCTATACCGTGTGGGGATCACCATGGATATATCGAACCGGAACC	600
Db	181	HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr	200
QY	601	TCCGTCAATTGTATAGTAGAAATGGATGCCGCTCTGTGTTTCGGTATTTCATATTTT	660
Db	201	SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe	220
QY	661	GCAATGCCCAATGCGGACATCGGAACATATCTCCATTTTATGGTCTATCCCCACACAGAG	720
Db	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu	240
QY	721	GCTGCCGAGAACCCATCGGATATCCCCAGGATAATTTCAAACAATAGATAGCTATTTT	780
Db	241	AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelLysGlnLeuAspSerTyrPhe	260
QY	781	TCAATGGATTGGACAAGCGTCGAAAAGCAGCGCTCCAGTCAAGCGTAACTTCTCATC	840
Db	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280
QY	841	ACATCACACTTTCACAGTTGGGTGGGACTGGGCTCCAAAACACTACTCTGTATGTTCAATG	900
Db	281	ThrSerHisPheThrValGlyTyrPaspTrpAlaProLysThrThrArgValCysSerMet	300
QY	901	ACTAAGTGGAAAGAGGTGACTGAATGTTCCGTGCAACAGTTAATGGGAGATACAGATTT	960
Db	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320
QY	961	ATGGCCCGTGAACCTTCGGCAACGTTTATCAGTAATACGACTGAGTTTGATCCAAATCGC	1020
Db	321	MetaLargGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340

QY	1021	ATCATATTAGGACAATGTATTAAACGCGAGCGAGACAGCAATTCGAGCAGATATTTAGG	1080
DB	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360
QY	1081	ACAAAATATATGACAGTCAAGTTCGAGCATGTACAAATTTCTTGGCTCTCGGG	1140
DB	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380
QY	1141	GGATTTATTCTAGCATATCAGCCTGTCTATCCAAATCCCTGGCTCATATGTACTCTAGA	1200
DB	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400
QY	1201	GAATTGATGAGAGACAAACAGACCGATGAGATGCTCGACCTGGTGAACAATAAGCATGCA	1260
DB	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420
QY	1261	ATTATTAAAGAAAATGCTACTCTCATTGTCCAGGATTCGGCGGAGATATTCGAAATGCACCA	1320
DB	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgAspIleArgAsnAlaPro	440
QY	1321	AATAGAAAAATAACATTAGACGACACACAGCTATTAAATCGACATCTGTCTCAATTC	1380
DB	441	AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe	460
QY	1381	GCCATGCTCCAATTTCTTTATGATCATATACAAACCATATTAATATGATATGTTTAGPAGG	1440
DB	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480
QY	1441	ATTGCCACAGCTGGTGGAAATTCGCAATAGAGAACTGTTTATGGCAGCAGGAGATA	1500
DB	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle	500
QY	1501	AAGATTAACTCTAGCGCTACAGCGAGTGCACATTTAGAAAGAGAGTGGCTGCAACAGATG	1560
DB	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet	520
QY	1561	TTGGGGGATGTCGCTGCTGTATCGAGCTGCACGTCTATAGATCGGGAATCCGTCACCTTG	1620
DB	521	LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540
QY	1621	CAAAATTTCTATGGAGTTATCACATCCACTCAATACATGTTATAGCCACCATTTGGTTCTA	1680
DB	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu	560
QY	1681	TTTTTCATATGGAGAAAACCAAGGAAACATACAGGAGCAACTCGGTGAAAAACAAGAGTTG	1740
DB	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu	580
QY	1741	CTTCCCAACGCTAGAGGCTGTAGCCCATGCTCGGCTAATCATCGTAGATATTTCTCTTTT	1800
DB	581	LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe	600
QY	1801	GGATCCCGTTATGCTTTTATTTGAAAACATAATTTCTTAAGATGGTAGACGTCGCCGAT	1860
DB	601	GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAsp	620
QY	1861	ATACAGATTGCTAGCACATTGTGAGCTTAATCTAACCCCTGTAGAGATCGGAAATTT	1920
DB	621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640
QY	1921	TTGCTTTTATCCGTTTACAAAAAGAGAGTTCGGTGTATGTTGGTATTTGGATTATGCA	1980
DB	641	LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla	660
QY	1981	GAAGTAGCTCGCGCAATCAACTACATGAATTTAAATTTTATGACATAAACAAGTAATA	2040
DB	661	GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle	680
QY	2041	GAAGTGGATACAAATTACGCG-----GGCTGCAGGAATTCGCGCTGTATGCTCTATT	2091
DB	681	GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu-----	695

QY	2092	ACTATAAAAGATGCAAAACCCAAATATAATATGGCCAAACCCAAATTA---GAAGCAGCGCCGAATG	2144
DB	696	-----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys	708
QY	2149	GAGTTAACAGATCTAATCAATGCTAAAGCGCATGACATTAGCTTCCACTACAAGACTATGCC	2208
DB	709	valValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet	728
QY	2209	AAG-----ATTGAAGCTAGTCTTTATCATCTGCT	2235
DB	729	SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleIleAlaGlyLeuValAlaAla	748
QY	2236	TATAGTGAAGCTGAACAGCTTAAC-----AATTAACCTTAATGCA-----	2274
DB	749	PheLeuAlaTyrArgTyrValAlaGlnLysLeuLysSerAsnProMetLysAlaLeuTyrPro	768
QY	2275	---ACATTAGACAACACTAAAATGGCTAAAACCTAATTTAGAAATCAGCCATCAACCAAGCT	2331
DB	769	MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp	788
QY	2332	ATATAGGATAAACGAGCTTTTGATATGAACACCCAAATTTAGTTGAAGCA-----	2382
DB	789	AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet	806
QY	2383	-----TACAAACGACTAAAACCCACTTCTAGAACACAGCTGCTACTACCTT	2427
DB	807	IleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu	823
RESULT 3			
US-08-213-449A-13			
; Sequence 13, Application US/08213449A			
; GENERAL INFORMATION:			
; APPLICANT: NAZERIAN, Keyvan			
; APPLICANT: LEE, Lucy F.			
; APPLICANT: YANAGIDA, Noboru			
; APPLICANT: OGAWA, Ryohei			
; APPLICANT: LI, Yi			
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR			
; PROTECTION AGAINST MAREK'S DISEASE			
; NUMBER OF SEQUENCES: 19			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH			
; STREET: P.O. Box 747			
; CITY: Falls Church			
; STATE: Virginia			
; COUNTRY: USA			
; ZIP: 22040-0747			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patentin Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/213,449A			
; FILING DATE: 15-MAR-1994			
; CLASSIFICATION: 424			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Murphy Jr., Gerald M.			
; REGISTRATION NUMBER: 28,977			
; REFERENCE/DOCKET NUMBER: 1644-108P			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (703) 205-8000			
; TELEFAX: (703) 205-8050			
; INFORMATION FOR SEQ ID NO: 13:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 865 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-213-449A-13			

Alignment Scores:
Pred. No.: 0 Length: 865

Score:	3601.00	Matches:	719
Percent Similarity:	87.81%	Conservative:	16
Best Local Similarity:	85.90%	Mismatches:	60
Query Match:	63.57%	Indels:	42
DB:	6	Gaps:	8
US-09-147-052-3 (1-3261) x US-08-213-449A-13 (1-865)			
QY	1	ATGCACCTATTTAGCGGAATTCATATATTTTCCCTATATAGTTATTTCTATATGTTACGAAAC	60
DB	1	Meth1stYrPheArqArgAsnCysIlePhePheLeuIleValIleLeuTyrglyThrAsn	20
QY	61	TCATCTCCGAGTACCCAAAATGTGCACATCAAGAGAAGTTGTTTCAGAGCGTCAGTTGCT	120
DB	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40
QY	121	GAGGAAGAGTCTACGTTTATCTTTGTCCCCCACCAGTGGGTTCAAACCGTGATCCGTCFA	180
DB	41	GluGluGluSerThrPheTyrlLeuCysProProValGlySerThrValIleArgLeu	60
QY	181	GAACGCCGCGAAATGTCCTCGAACCTAGAAAAGCCACCGAGTGGGTGCAAGAAATCGCG	240
DB	61	GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla	80
QY	241	ATATATTTAAAGAAATATCAGTCCATATAAATTTAAAGTGACCGCTTTATTTATAAAAT	300
DB	81	IleLeuPheLysGluAsnIleSerProTyrlLysPheLysValThrLeuTyrrLysAsn	100
QY	301	ATCATTCAGACGACGACATGACGGGGAGCACATATAGACAGATCACTAATCGATATACA	360
DB	101	IleIleGlnThrThrTrpThrGlyThrThrTyrrArgGlnIleThrAsnArgTyrrThr	120
QY	361	GATAGGACGCCGTTTCCATTGAAGAGATCACGGATCTAATCGACGGCAAGAAAGATGCG	420
DB	121	AspArgThrProValSerIleGluGlnIleThrAspLeuIleAspGlyLysGlyArgCys	140
QY	421	TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGACAGGATGCG	480
DB	141	SerSerLysAlaArgTyrlLeuArgAsnAsnValTyrrValGluAlaPheAspArgAspAla	160
QY	481	GGAGAAAAACAAGTACTTCTAAACCATCAAAATCAACACGCCCGGAATCTAGGCGATCG	540
DB	161	GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180
QY	541	CACACGACTAATGAGACGTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACC	600
DB	181	HisThrThrAsnGluThrTyrrThrValTrpGlySerProTrpIleTyrrArgThrGlyThr	200
QY	601	TCCGTCATTTGTATAGTAGAGGAATGTGATGCCCGCTCTGTGTTTCCGTTATTCATATTTT	660
DB	201	SerValAsnCysIleValIleGluMetAspAlaArgSerValPheProTyrrSerTyrrPhe	220
QY	661	GCAATGGCAATGGCGACATCGCGAAACATATCTCCATTTTATGGTCTATCCGCCACACAG	720
DB	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrglyLeuSerProProGlu	240
QY	721	GTCGCCGAGAACCCATCGGATATCCCCAGGATAATTTCAACAACACTAGATAGCTATTTT	780
DB	241	AlaAlaAlaGluProMetGlyTyrrProGlnAspAsnPheLysGlnLeuAspSerTyrrPhe	260
QY	781	TCATGGATTTGGACAAGCGTCGAAAGCAAGCGTTCCAGTCAAGCGTAACTTTCTCATC	840
DB	261	SerMetAspLeuAspLysArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280
QY	841	ACATCACACTTCACAGTTGGTGGGACGTGGGCTCCAAAAACTACTCGTGTATGTTCAATG	900
DB	281	ThrSerHisPheThrValIglyTrpAspTrpAlaProLysThrThrArgValCysSerMet	300
QY	901	ACTAAGTGGAAAGAGGTGCACGTGAATGTTGCGTGCACACATTAATGGGAGATACAGATTT	960
DB	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrrArgPhe	320
QY	961	ATGCCCGGTGAACCTTTCGGCAACGTTTATCAGTAATACGACTGAGTTTGATCCAAATCGC	1020

QY 961 ATGGCCCGTGAACTTTCGGCAACGTTTATCAGTAATACGACTGAGTTTGATCCAAATCGC 1020

Alignment Scores:

Pred. No.:	0	Length:	865
Score:	3601.00	Matches:	719
Percent Similarity:	87.81%	Conservative:	16
Best Local Similarity:	85.90%	Mismatches:	60
Query Match:	63.57%	Indels:	42
DB:	6	Gaps:	8

US-09-147-052-3 (1-3261) x US-08-213-449B-13 (1-865)

Qy	1	ATGCAC	TACTATTTAGCGG	AATTCG	CAATATTTTCC	TTATAG	TTATTTCTATATAG	TTATTTCTATATAG	TTACG	TAAC	60			
Db	1	Meth1	stYrPheA	rgAsnCys	ilePheLeu	ileValle	ValleLeu	YrGlyThr	Asn	20				
Qy	61	TCAT	CTCGAG	TACCCAAA	TGTGC	ACATCA	GAGAA	GTGTTTC	GAGCG	TCCAG	TTGCT	120		
Db	21	SerSer	ProSer	ThrGln	AsnVal	ThrSer	ArgGlu	ValVal	SerSer	ValGln	LeuSer	40		
Qy	121	GAGG	AAAGAG	TCTACG	TTTATCT	TTGTCT	CCCCC	ACAC	GAGTGG	GTGTCA	ACCGTGTAC	CGCTCTA	180	
Db	41	GluGlu	GluSer	ThrPhe	YrLeu	CysPro	ProVal	GlySer	ThrVal	ileArg	Leu	60		
Qy	181	GAAC	CGCGCG	GAANA	TGTC	CGAAC	CTAGA	AAAGCC	ACGAGT	GGGTGG	GGAAG	ATCGG	240	
Db	61	GluPro	ProArg	LysCys	ProGlu	ProArg	LysAla	ThrGlu	TrpGly	CluGly	ileAla	80		
Qy	241	ATAT	TATTTAA	GAGA	AAATCAG	TCCAT	ATATA	AAATTA	AGTCA	CGCTTAT	TATATA	AAAT	300	
Db	81	IleLeu	PheLys	GluAsn	ileSer	ProYr	LysPhe	LysVal	ThrLeu	YrYr	LysAsn	100		
Qy	301	ATCAT	TCAG	ACGAC	ATGAC	GGGG	CGAC	ATATAG	ACAC	ATCACT	AAATCG	ATACA	360	
Db	101	Ileile	GlnThr	ThrThr	TrpThr	GlyThr	ThrTyr	ArgGln	ileThr	AsnArg	YrThr	120		
Qy	361	CATAG	GAGCG	CGCTTC	CATTG	AGAG	ATCAG	GATCTA	ATCG	AGCGG	CAAGAG	ATCG	420	
Db	121	Aspar	thrThr	ProVal	SerIle	GluGlu	ileThr	AspLeu	ileAsp	GlyLys	GlyArg	Cys	140	
Qy	421	TCAT	CTAA	GCAAG	ATACCT	TAGAA	CAATGT	ATGTT	GAAG	CGTTTC	CACAG	GATCGG	480	
Db	141	SerSer	LysAla	ArgTyr	LeuArg	AsnVal	TyrVal	GluAla	PheAsp	ArgAsp	Ala	160		
Qy	481	GGAG	AAAA	CAAGT	ACTTCT	TAAAA	CCATCA	AAATTTCA	ACAC	CGCCG	GAATCT	TAGGCG	540	
Db	161	GlyGlu	LysGln	ValLeu	LeuLys	ProSer	LysPhe	AsnThr	ProGlu	SerArg	AlaTrp	180		
Qy	541	CACAG	GACTA	TGAG	ACGTAT	ACCGT	GTGGG	ATCACC	ATGG	ATATCG	ACAGCG	GAAC	600	
Db	181	H1st	ThrThr	AsnGlu	ThrTyr	ThrVal	TrpGly	SerPro	TrpIle	YrArg	ThrGly	Thr	200	
Qy	601	TCCG	TCAAT	TGTAT	AGAG	GAATG	ATGCC	CGCTGT	GCTG	TTTCCG	TATTCAT	ATATTT	660	
Db	201	SerVal	AsnCys	ileVal	IGlu	MetAsp	AlaArg	SerVal	PhePro	YrSer	YrPhe	220		
Qy	661	GCAAT	GGCCA	TGCG	GACATCG	CGAAC	ATATCT	CCATTTTAT	TGGTCT	ATATCC	CCAC	CAGAG	720	
Db	221	AlaMet	AlaAsn	GlyAsp	ileAla	Asnile	SerPro	PheYr	GlyLeu	SerPro	ProGlu	240		
Qy	721	GCTG	CGCG	GAGAC	CCCAT	CGGAT	ATATCC	CCAGG	ATAATTTCA	AAAC	ACTAG	ATAGCTAT	780	
Db	241	AlaAla	AlaGlu	ProMet	GlyTyr	ProGln	AspAsn	PheLys	GlnLeu	AspSer	YrPhe	260		
Qy	781	TCAAT	GGA	TTTGG	CAAC	GCTG	GAAG	CAAGC	CTTCC	CAGTCA	ACGGTAA	CTTCTCAT	840	
Db	261	SerMet	AspLeu	AspLys	ArgGly	SerVal	SerLeu	ProVal	LlyArg	AsnPhe	Leuile	280		
Qy	841	ACAT	CAC	ACTT	CAC	AGTTGG	GCTGG	GACTCG	GGGTCC	CAAAA	ACTACT	CTGTA	TGTTCAA	900
Db	281	ThrSer	H1stPhe	ThrVal	GlyTrp	AspTrp	AlaPro	LysThr	ThrArg	ValLys	SerMet	300		
Qy	901	ACTA	AGTGG	AAAG	AGGTG	ACTG	AAATG	TGTCG	TGCA	CGATTA	TGTTAA	TGTTGG	GAGATAC	960
Db	301	ThrLys	TrpLys	GluVal	ThrGlu	MetLeu	ArgAla	ThrVal	AsnGly	ArgThr	ArgPhe	320		

Qy	961	ATGCCCGTGAAC	TTTCGGCAACG	TTTATCAGTAATACGACTGAGTTGATCCAAATCGC	1020
Db	321	MetaAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340		
Qy	1021	ATCATATTAGGACAAATG	TATTAAACGGCAGCGCAGACCAATTCGAGCAGAGATATTTAGG	1080	
Db	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360		
Qy	1081	ACAAATATATATGACAGCTCAGCTCAAGGTTGGACATGTACAATATTTC	TTCGTGGCTCGCGG	1140	
Db	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380		
Qy	1141	GGATTTATTAGCATATCAGCCCTGTTCTATCCAAATCCCTGGCTCATATGTACCTCAGA	1200		
Db	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400		
Qy	1201	GAATTGATGAGAGACAACAGCAGCGGATGAGATGCTCGACCTGGTAAACAATATAGCATGCA	1260		
Db	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420		
Qy	1261	ATTTATTAAGAAAATGCTACTCTATTCTCAGATTGGCGGAGATATTCGAATTCGACCA	1320		
Db	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440		
Qy	1321	AATAGAAAAATTAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTTCAATTC	1380		
Db	441	AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe	460		
Qy	1381	GCCATGCTCCAAATTC	TTTATGATCATATACAACCCCATATTAAATGATATGTTTAGTATGAGG	1440	
Db	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480		
Qy	1441	ATTGCCACAGCTTGGTCGAATTCGCAAAATGAGAGAACTCTGTTATGGCAGCAGAGGATA	1500		
Db	481	IleAlaThrAlaThrPcysGluLeuGlnAsnArgGluLeuValLeuThrPHisGluGlyIle	500		
Qy	1501	AAGATTAAATCCTACGCGTACAGCGAGTGCACACATTAGGAAGAGAGCTGGCTGCAAAAGATG	1560		
Db	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet	520		
Qy	1561	TTGGGGGATGTCGCTGCTGATCGAGCTGCACATGCTATAGATCGGGAATCCGTCACATTCG	1620		
Db	521	LeuGlyAspValAlaAlaValIleSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540		
Qy	1621	CAAAATTCATGCGAGTTATCACATCCACATTAATACATGTTATACCGCACCATTCGTTCTA	1680		
Db	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu	560		
Qy	1681	TTTTCATATGGAGAAAACCAAGGAAACATACAGGCAACTCGGTGAAAAACAACAGAGTTCG	1740		
Db	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlnLeuGlyGluAsnAsnGlnLeu	580		
Qy	1741	CTTCCAACGCTAGAGCGCTGATAGGCCATGCTCGGCTAATCATCGTAGATATTTCTGTTT	1800		
Db	581	LeuProThrLeuGluAlaValAluProCysSerAlaAsnHisArgArgTyrPheLeuPhe	600		
Qy	1801	GGATCCGGTTATGCTTTATTGAAAACTATAATTTTCTTAAGATGGTAGACGTCGCCGAT	1860		
Db	601	GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp	620		
Qy	1861	ATACAGATTGCTACACATTTGTCGAGCTTAATCTAACCCCTGCTPAGAAGATCGGGAATTC	1920		
Db	621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640		
Qy	1921	TTGCTTTTATCCGTTTACAAAAAGAGAGTTGCGGTGATCTTGCTGTATTTGATTTATGCA	1980		
Db	641	LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla	660		
Qy	1981	GAAGTAGCTCGCCGCAATCACTACATGAACTTAATTTTATGACATAAACAAAGTAATA	2040		
Db	661	GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle	680		


```

QY 2041 GAAGTGATACAAATTACGG-----GGGCTGCAGGAATTCGGCTGTATGCTCTATT 2091
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
681 GluValAspThrSerIleThrAlaPheMetAsnGlyLeuAlaGluLeu----- 695
QY 2092 ACTAAAAAAGATGCAAAACCCCAATAATATGGCCAAACCCCAATTA---GAAGCAGCGCGAATG 2148
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
696 -----PheAsnGlyMetClyGlnValGlyGlnAlaIleGlyLys 708
QY 2149 GAGTTACACATCAATCAATGCTAAAGCGATGACATAGCTTACACACAGACTATGCCC 2208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
709 ValValValGlyAlaAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
QY 2209 AAG-----ATTGAAGCTAGTTTATCATCTGCT 2235
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
QY 2236 TATAGTGAAGCTGAACACAGTTAAC-----AATAACCTTATATGCA----- 2274
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
749 PheLeuAlaIleValSerThrIleValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
QY 2275 ---ACATTAGAACACTAAAATGGCTAAACTAATTTAGAACTCAGCCATCAACCAAGCT 2331
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluGluSerAsp 788
QY 2332 AATACGGATAAAGCGACTTTTGATAATGAACACCCCAAAATTTAGTTGAAGCA----- 2382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet 806
QY 2383 -----TACAAACACTAAACACACTTTTAGAACACAGCTGCTACTAACCTT 2427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 IleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu 823

RESULT 5
US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Alignment Scores:
Pred. No.: 1-32e-172 Length: 456
Score: 2015.00 Matches: 396
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.57% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-3 (1-3261) x US-09-147-052-2 (1-456)
QY 2071 GAATTCGGCTGATGCTATTACTAAAAAGATGCAAAACCCCAATAATATGGCCAAACCCCAA 2130
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
QY 2131 TTAGAGCAGCGGATGAGTTACAGATCTAATCAATGCTTAACGGATGACATAGCT 2190
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100

```

```

QY 2191 TCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 2250
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
QY 2251 ACAGTTTAAACAATTAACCTTAATGCAACATTAACAACAATAAATAATGGCTAAACCTTAATTTA 2310
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
QY 2311 GAATCAGCCCATCAACCAAGCTTAATACGATTAACAAACGACTTTTGTATATGACACCCCAAT 2370
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
QY 2371 TTAGTTGAAGCATACAAGCAGCACTAAAAACACACTTTTAGAACAACTGCTACTAACCCTTGAA 2430
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
QY 2431 GCTTTGTCATCAACTGCTTAATTAATCAAAATTCGCAATAATTTAGTGGATCTATACAATAAA 2490
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
QY 2491 GCTAGTAGTTTAATACTAAACACTAGATCCACTTAATGGGGAAGCGCTTTTAGATTCT 2550
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
QY 2551 AATGAGATTACTACAGCTAATAAGAATTAATAATACGTTTATCAACTATTAATTAACACAA 2610
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
221 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
QY 2611 AAGACTAATGCTGATGCATCTATCTAATAGTTTATTAATAAAGTGATTCAAAATATGAA 2670
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 260
QY 2671 CAAAGTTTTGTAGGAGACTTTTACAAACGCTAATAGTTCAACCTTCAAACTACAGTTTGT 2730
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280
QY 2731 GCTTTTAGTCTGATGTACACACCCGTCATTAATAATAATATCAAGAAGCCGTTTGGAAAT 2790
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTyrAsn 300
QY 2791 GGTGATGAACCTTCAAGTAGAATCTTGCACCAACCAAGTAATAGTATCAGATGTTCTTGG 2850
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyr 320
QY 2851 ATTTATAGTTTACGTCGAACAAACACGAGTACCAATTTAGTTTACCAACTATGTTCCA 2910
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
QY 2911 TCAACTGGTTATTATTATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTT 2970
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
QY 2971 GGATTACAATACAAATTAATGGAATGTTTCAACAGTTTGCAGTTTGCACCTTCAACT 3030
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
361 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
QY 3031 AGTGCAAAATAACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGATGTTGCTAAA 3090
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
QY 3091 ATCGTTTTATCAGTTTAAAGATTGGCCAAACACAAATCGAATTAAGTTGTTCAACGGGT 3150
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
QY 3151 GAAGGAATATGAATAAAGTTGCGCCAATGATTGGCAACATTTATCTTACCTCAAAATGAA 3210
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
QY 3211 AATAATGCTGATAAGATCCCCGGGTACCGTCGACCGGTACATTTTTTA 3258
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

```



```
RESULT 6
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTW-1 portion of pNZ40K-S
US-09-901-572A-3
Alignment Scores:
Pred. No.: 5,66e-172 Length: 456
Score: 2008.00 Matches: 394
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 35.45% Indels: 0
DB: 24 Gaps: 0

US-09-147-052-3 (1-3261) x US-09-901-572A-3 (1-456)
QY 2071 GAATTCGGCTGATCTATTACTAAAGGATGCAACCCCAAAATTAATGGCCAAACCCAA 2130
DB 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnGlnThrGln 80
QY 2131 TTAGAAGCAGCGCGAATGGAGTAAACAGATCTCAATCAATGCTAAAGCGATGACATTAGCT 2190
DB 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
QY 2191 TCACTACAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAA 2250
DB 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
QY 2251 ACAGTTAACTAACTTAATCAATCAATAGCAACCTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2310
DB 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
QY 2311 GAATCAGGCATCAACCAAGCTAATACGAGTAAAGAGCTTTTGATATGAACACCCCAAT 2370
DB 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
QY 2371 TTAGTTGAAGCATCAACAGCCTAAACCCACTTTAGAACACGCTGCTACTAACCTTGA 2430
DB 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
QY 2431 GGTTCGTATCACTCACTTATTAATCAATTCGCAATTAATTTAGTGGATCTATACAAATAA 2490
DB 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
QY 2491 GCTAGTATTTTAATAACATAACACATAGATCCCAATTAATGGGGGACGCTTTTATGATCT 2550
DB 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
QY 2551 AATGAGATTACTACAGCTAATTAAGAATTAATAATATAGCTTATCACTTATTAATGAACAA 2610
DB 221 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
QY 2611 AAGACTAATGCTGATGCTATCTAATTAATAGTTTATTAATAAAGTGAATCAAAATATGA 2670
DB 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 260
QY 2671 CAAAGTTTGTAGGAGCTTTTCAACACGCTAATGTTCAACCTTCAACCTACAGTTTGT 2730
DB 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280

QY 2731 GCTTTTAGTCTGATGTAACACCCGCTCAATTATAAATATGCAAGAGGACCGTTTGGAAAT 2790
DB 281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgArgThrValTrpAsn 300
QY 2791 GGTGATGAACCTTCAAGTAGAATCTTGCAACACAGCAATAGTATACAGATGTTCTCTGG 2850
DB 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 320
QY 2851 ATTTATAGTTTACCTGGGAACACAGAGTACCAATTTAGTTTATAGCAACATATGCTCCA 2910
DB 321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
QY 2911 TCAACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTTAATACGTT 2970
DB 341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
QY 2971 GGATTACAATAACAAATTAATTAATGAAATGTTCAACAGTTGAGTTTCCCACTTCAACT 3030
DB 361 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
QY 3031 AGTGCATAATATACACTACAGCTAATCCAACTCCAGCAGTTGTAGTACAGATTAAAGTTGCTAAA 3090
DB 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
QY 3091 ATCGTTTATATCAGTTTAAAGATTGGCCAAAACACAACTCAATTAAGTGTCCACGGGT 3150
DB 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
QY 3151 GAAGGAAATATCAATAAAGTTGCGCAATGATTGGCAACATTATCTTAGCTCAAAATGAA 3210
DB 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
QY 3211 AATAATGCTGATTAAGATCCCGGTTACCGTCCGACCCGCTACATTTTATA 3258
DB 441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

RESULT 7
US-10-131-591A-12
; Sequence 12, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTW-1 portion of pNZ40K-S
US-10-131-591A-12
Alignment Scores:
Pred. No.: 6,73e-165 Length: 384
Score: 1929.50 Matches: 382
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.22% Mismatches: 1
Query Match: 34.06% Indels: 1
DB: 27 Gaps: 1

US-09-147-052-3 (1-3261) x US-10-131-591A-12 (1-384)
QY 2071 GAATTCGGCTGATGCTATTACTAAAGGATGCAACCCCAAAATTAATGGCCAAACCCAA 2130
DB 1 GluPhe---CysMetSerIleThrLysLysAspAlaAsnProAsnGlnThrGln 19
QY 2131 TTAGAAGCAGCGCGAATGGAGTAAACAGATCTCAATCAATGCTAAAGCGATGACATTAGCT 2190
DB 20 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 39
```

QY 2191 TCACTACAAGACTATGCGAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 2250
 Db 40 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerAlaTyrSerGluAlaGlu 59
 QY 2251 ACAGTTAACAAATACCTTAATGCAACATTAGAACAATAAAATGGCTAAACCTAAATTTA 2310
 Db 60 ThrValAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 79
 QY 2311 GATCAGCCATCAACCAAGCTAATACGGATAAAGAGCTTTTGTATATGACACCCCAAT 2370
 Db 80 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 99
 QY 2371 TTAGTTGAAGCATACAAAGCACTAAACCACTTTAGAACAACTGCTACTAACCTTGAA 2430
 Db 100 LeuValGluAlaTyrLysAlaLeuLysThrLeuGluGlnArgAlaThrAsnLeuGlu 119
 QY 2431 GGTTCATCATCACTGCTTAATCAAAATTCGCAATTAATTTAGTGGATCTATACAAATA 2490
 Db 120 GlyLeuSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 139
 QY 2491 GCTAGTAGTTTAATACTAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCT 2550
 Db 140 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 159
 QY 2551 AATGAGATTACTACAGCTAATAAGAAATTAATAATACGTTATCAACTATTAAATGAACA 2610
 Db 160 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 179
 QY 2611 AAGACTAATGCTGATGATTAATCAATGATTTTATTAATAAAGTGAATCAAAATAATGAA 2670
 Db 180 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 199
 QY 2671 CAAAGTTTCTAGGACTTTTACAACGCTTAATGTTCAACCTTCAAACTACAGTTTGT 2730
 Db 200 GluSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 219
 QY 2731 GCTTTTAGTCTGATGTAACACCGCTCAATTAATAATATGCAAGAGCGCTTTTGAAT 2790
 Db 220 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTrpAsn 239
 QY 2791 GGTGATGACCTTCAGTAGAATCTTGCNAACAGAAATAGTATCAGATGTTCTTGG 2850
 Db 240 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp 259
 QY 2851 ATTTATAGTTTACTGGACAAACACGAGTACCAATTTAGTTTACCACTATGCTCCA 2910
 Db 260 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 279
 QY 2911 TCAACTGGTTATTTATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTT 2970
 Db 280 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 299
 QY 2971 GGATTACAATAAATAATATGGAATGTTCAACAGTTGAGTTTGGCCTTCAACT 3030
 Db 300 GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 319
 QY 3031 AGTGCAAAATAACTACAGCTAATCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAA 3090
 Db 320 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 339
 QY 3091 ATCGTTTATTCAGGTTTAAAGATTGCGCCAAACACAAATCGAATTAAGTTGTTCAACGGGT 3150
 Db 340 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 359
 QY 3151 GAAGGAATAATGATTAAGTTGGCCCAATGATTGGCAACATTTATCTAGCTCAAAATGAA 3210
 Db 360 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 379
 QY 3211 AATAATGCTGATAAG 3225
 Db 380 AsnAsnAlaAspLys 384

RESULT 8

US-08-804-439-21
 ; Sequence 21, Application US/08804439
 ; GENERAL INFORMATION:
 ; APPLICANT: Rose, Timothy M.
 ; APPLICANT: Bosch, Marnix L.
 ; APPLICANT: Strand, Kurt
 ; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morrison & Foerster
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,439
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schiff, J. Michael
 ; REGISTRATION NUMBER: 40,253
 ; REFERENCE/DOCKET NUMBER: 29938-20002.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 868 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-804-439-21
 Alignment Scores:
 Pred. No.: 5,31e-151 Length: 868
 Score: 1777.50 Matches: 375
 Percent Similarity: 61.30% Conservative: 143
 Best Local Similarity: 44.38% Mismatches: 259
 Query Match: 31.38% Indels: 69
 DB: 12 Gaps: 16

US-09-147-052-3 (1-3261) x US-08-804-439-21 (1-868)
 QY 91 AGAAGAGTTGTTTCGAGCGTCCAGTTGCTGAGAGAGAGTCTACGTTTATCTTTGTCCTC 150
 Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60
 QY 151 CCACAGTGGTTCACCGTATCCGCTAGAACCCGCGCGGAAATGTCGCCAGACCTAGA 210
 Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80
 QY 211 AAAGCCACCGAGTGGGTGAGGAATCGCGATATTTAAGAGAGAAATATCATGTCATAT 270
 Db 81 LeuGlyLysAsnPheThrGluGlyIleAlaValTyrLysGluAsnIleAlaLys 100
 QY 271 AAATTTAAAGTGACGCTTTTATTATAAAATATCATTCAGACGACGACATGGAGCGGAGC 330
 Db 101 LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrIleAlaIleAlaLys 120
 QY 331 ACATATAGACAGATCACTAATCGATATACAGTAGACCGCGCTTTCCATTGAAGAGATC 390
 Db 121 SerTyrThrGlnIleThrAsnArgTyrAlaAspValProIleProValSerGluIle 140
 QY 391 ACGGATCTAATCGACGGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAACAAAT 450

```

141 ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn 160
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 GTATATGTTGAACGGTTTACAGGGATCGGGAGAGAAACAAAGTACTTAAACCATCA 510
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
511 AAATTCACACGCCGAATCTAGGCATGGCACAGCACTAAGAGACGTATACCGTGG 570
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 LysTyrAsnSerValGlySerLysAlaIlePheHisThrThrAsnAspThrTyrMetValAla 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
571 GGATCACCATGGATATATCAACGGGAACCTCCGTCATTTGGCAATGTATAGTAGGAATGGAT 630
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluValGlu 220
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 GCCCGCTGCTGTTCCGTTATTCATATTTGGTCAATTTGGCCAAATGGGACATCGCAACATA 690
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSerThrGlyAspIleIleTyrMet 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 TCTCCATTTATGGTCTATCCCAACAGAGGCTCGCGAGAACCCATGGATATCCCCAG 750
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 SerProPheGlyLeuArg---AspGlyAlaTyrArgGluHisSerAsnTyrAlaMet 259
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 GATAATTTCAACAACATAGATAGTATTTTCAATGGATTGGACAAGCGTCGAAAAGCA 810
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 AspArgPheHisGlnPheGluGlyTyrArgGlnArgAspLeuAspThrArgAlaLeuLeu 279
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
811 AGCCTTCCAGTCAACGGTAACTTCTCATCATCACACTTCACAGTTGGGGGACTGG 870
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 Glu---ProAlaAlaArgAsnPheLeuValThrProHisLeuThrValGlyTrpAsnTrp 298
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
871 GCTCCAAAACACTACTCGTGTATGTTCAATGACTAAGTGAAGAGGTGACTGAATGTG 930
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LysProLysArgThrGluValCysSerLeuValLysTrpArgGluValGluAspValVal 318
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
931 CGTGCAACAGTTAAATGGGAGATACAGATTTATGGCCCGTGAATCTTCGCCAACGTTTATC 990
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 ArgAspGluTyrAlaHisAsnPheArgPheThrMetLysThrLeuSerThrThrPheIle 338
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 AGTAATACAGCTGAGTTTGATCCAAATCGCATCATATTAGGCAATGTATTAAACCGCAG 1050
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 SerGluThrAsnGluPheAsnLeuAsnGlnIleHisLeuSerGlnCysValLysGluGlu 358
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1051 GCAGAGCAGCAATCGCAGCAGATATTAGGACAAATATATAAGTACAGTCAAGTT 1110
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 AlaArgAlaIleIleAsnArgIleTyrThrArgTyrAsnSerSerHisValArgThr 378
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1111 GGACATGTACAATATTTCTGCTCTCGGGGATTTATGTAGCATATCAGCTGTCTTA 1170
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 GlyAspIleGlnThrTyrLeuAlaArgGlyGlyPheValValPheGlnProLeuLeu 398
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1171 TCCAAATCCCTGGCTCATATGTAACCTCAGAGAATTGATGAGACACAAACAGCCGATGAG 1230
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 SerAsnSerLeuAlaArgLeuTyrLeuGlnGluLeuValArgGlu----- 413
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1231 ATGCTCGACCTGGTAAACAAATAGCATGCAATTTATAGAAAATGCTACCTCATGTCA 1290
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 -----AsnThrAsnHisSer---ProGlnLysHisProThrArgAsnThr 427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1291 CGATTGCGCGGAGATATTCGAAATGCACCAATAGAAAATTAACATTAGACGACACCA 1350
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
428 ArgSerArgArgSerVal-----ProValGluLeuArgAlaAsnArg 441
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1351 GCTATTAAATCGACATCGTCTGTTCAATTCGCCATGCTCCAATTTCTTATGATCATATA 1410
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
442 ThrIleThrThrThrSerSerValGluPheAlaMetLeuGlnPheThrTyrAspHisIle 461
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1411 CAAACCCATTAATGATATGTTTAGTAGGATGGCACAGCTGGTGGCAATGCGAAT 1470
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 GlnGluHisValAsnGluMetLeuAlaArgIleSerSerSerTrpCysGlnLeuGlnAsn 481
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1471 AGAGAACTGTTTATGGCACGAGGATAAAGATTAATCCTAGCGCTACAGCGAGTGCA 1530
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 2514 ACTAGATCCACTA 2526
 Db 845 gLeuthrGlyLeu 849

RESULT 9
 US-09-301-390-21
 : Sequence 21, Application US/09301390
 : GENERAL INFORMATION:
 : APPLICANT: Rose, Timothy M.
 : APPLICANT: Bosch, Marnix L.
 : APPLICANT: Strand, Kurt
 : TITLE OF INVENTION: GLYCOPROTEIN B OF THE REFV/KSHV
 : TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
 : NUMBER OF SEQUENCES: 100
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Morrison & Foerster
 : STREET: 755 Page Mill Road
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304-1018
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC Compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/720,229
 : FILING DATE: 26-SEP-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Schliff, J. Michael
 : REGISTRATION NUMBER: 40,253
 : REFERENCE/DOCKET NUMBER: 29938-20002.00
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 813-5600
 : TELEFAX: (415) 494-0792
 : TELEX: 706141
 : INFORMATION FOR SEQ ID NO: 21:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 868 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-09-301-390-21

Alignment Scores:
 Pred. No.: 5,31e-151 Length: 868
 Score: 1777.50 Matches: 375
 Percent Similarity: 61.30% Conservative: 143
 Best Local Similarity: 44.38% Mismatches: 259
 Query Match: 31.38% Indels: 69
 DB: 17 Gaps: 16

US-09-147-052-3 (1-3261) x US-09-301-390-21 (1-868)

QY 91 AGAGAAAGTTGTTTCGAGCGTCCAGTTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCC 150
 Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheThrValCysPro 60

QY 151 CCACAGTGGTTCACACCGTGATCCGCTAGAACCCGCGCAAAATGTCGCCAGCACTAGA 210
 Db 61 ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis 80

QY 211 AAGCCACCCAGTGGGGTGAAGGAATCCGCATATTAATTAAGAGAAATATCATGTCATAT 270
 Db 81 LeuGlyLysAsnPheThrGluGlyIleAlaValValThrLysGluAsnIleAlaAlaTyr 100

QY 271 AAATTTAAAGTGACGCTTTTATTATAAAATATCATTCAGACGAGCATGTGACGGGACG 330
 Db 101 LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGlySer 120

QY 331 ACATATAGACAGATCACTAATCGATATACATAGGAGCGCCGTTTCCATTTAGAGATC 390
 Db 121 SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle 140

QY 391 ACGGATCTAATCGACGGCAAGGAGATGCTCATCTAAAGCAAGCATACCTTTAGAACAT 450
 Db 141 ThrAspThrIleAspLysPheGlyLysCysSerLysAlaThrTyrValArgAsn 160

QY 451 GTATATGTTGAAGCGTTTGCACAGGATCGCGGAGAAAAACAAAGTACTCTCTAAAACCA 510
 Db 161 HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer 180

QY 511 AAATTCACACGCGCGAATCTAGGCGATGGCACAGCACTAATGAGCATATACCGTGTGG 570
 Db 181 LysTyrAsnSerValGlySerLysAlaTrpHisThrAsnAspThrTyrMetValala 200

QY 571 GGATCACCATGGATATATCGAACGCGAACCCTCCGCTCAATTCGTATAGTAGAGGAATGGAT 630
 Db 201 GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu 220

QY 631 GCCGCTCTGTGTTCCGTATTCATATTTGCAATGGCCCAATGCGCATCGCGCAACATA 690
 Db 221 AlaArgSerIlePheProTyrAspSerPheGlyLeuSerThrGlyAspIleIleTyrMet 240

QY 691 TCTCCATTGTTGCTCTATCCCCACAGAGCGTCCCGCAGAACCCATGGGATGCCCGAG 750
 Db 241 SerProPheGlyLeuArg---AspGlyAlaTyrArgGluHisSerAsnTyrAlaMet 259

QY 751 GATAATTTCAACAACTAGATAGCTATTTTCAATGGATTGGACAAGCTCGCAAAAGCA 810
 Db 260 AspArgPheHisGlnPheGluGlyTyrArgGlnArgAspLeuAspThrArgAlaLeuLeu 279

QY 811 AGCTTCCAGTCAAGCGTAATCTCTCATCACATCACATTCACATCGAGTTGGTGGGACTGG 870
 Db 280 Glu---ProAlaAlaArgAsnPheLeuValThrProHisLeuThrValGlyTyrAsnTrp 298

QY 871 GCTCCAAAACTACTCGTGTATGTTCAATGAGTGAAGTGAAGAGGTGACTGAATGTTG 930
 Db 299 LysProLysArgThrGluValCysSerLeuValLysTrpArgGluValGluAspValVal 318

QY 931 CGTGCAACAGTTAATGGGAGATACAGATTTATGGCCGCTGCACTTTTCGCAACGTTTATC 990
 Db 319 ArgAspGluTyrAlaHisAsnPheArgPheThrMetLysThrLeuSerThrThrPheIle 338

QY 991 AGTAATACGACTGAGTTTGATCCCAATCCGATCATATTAGGACAATGATTAACGCGAG 1050
 Db 339 SerGluThrAsnGluPheAsnLeuAsnGlnIleHisLeuSerGlnCysValLysGluGlu 358

QY 1051 GCAGAAAGCAATCGAGCAGATATTTAGGACAAAATATATACATCACGTCACGTCAGGTT 1110
 Db 359 AlaArgAlaIleIleAsnArgIleTyrThrArgTyrAsnSerSerHisValArgThr 378

QY 1111 GGACATGTCAATATTTCTGCTCTCGGGGATTTATTGTAGCATATCAGCTGCTCTA 1170
 Db 379 GlyAspIleGlnThrTyrLeuAlaArgGlyGlyPheValValPheGlnProLeuLeu 398

QY 1171 TCCAAATCCCTGCTCATATGTACCTCAGAGATTTGATGAGAGACAACAGGACCGATGAG 1230
 Db 399 SerAsnSerLeuAlaArgLeuTyrLeuGlnGluValArgGlu----- 413

QY 1231 ATGCTCGACTGGTAAACAAATAGCATGCAATTTATAAGAAAAATGCTACCTCATTTGTC 1290
 Db 414 -----AsnThrAsnHisSer---ProGlnLysHisProThrArgAsnThr 427

QY 1291 CGATTCGCGGAGATATTCGAAATGCACCACCAATAGAAAAATATACATAGACGACACCA 1350
 Db 428 ArgSerArgArgSerVal-----ProValGluLeuArgAlaAsnArg 441

QY 1351 GCTATTAAATCGACATCGCTGTTCAATTCGCCCATGCTCCCAATTTCTTTATGATCATATA 1410

```

Db      442 ThrIleThrThrThrSerValGluPheAlaMetLeuGlnPheThrThrAspHisIle 461
QY      1411 CAAACCCATTAATGATGTTTGTAGTAGTGGCAGAGTGTGGCAATGCGAAT 1470
Db      462 GlnGluHisValAsnGlnMetLeuAlaArgIleSerSerSerTipCysGlnLeuGlnAsn 481
QY      1471 AGAGAACTGTTTATGGCAGCAGGGAATAGATTAATCCTAGCGCTACACGGAGTGA 1530
Db      482 ArgGluArgAlaLeuThrSerGlyLeuPheProIleAsnProSerAlaLeuAlaSerThr 501
QY      1531 ACATTAGGAAGAGAGTGGCTCAAGATGTTGGGGATGTCGCTGATATGAGCTGC 1590
Db      502 IleLeuAspGlnArgValIysAlaArgIleLeuGlyAspValIleSerValSerAsnCys 521
QY      1591 ACTGCTATAGATCGGAATCC---GTCACCTTGGCAAAATCTATCGGATGATCATCATCC 1647
Db      522 ProGluLeuGlySerAspThrArgIleLeuGlnAsnSerMetArgValSerGlySer 541
QY      1648 ACTAATACATGTTATAGCGGACCATGTTCTATTTTCATATGGAGAAACCAAGGAAC 1707
Db      542 ThrThrArgCysTyrSerArgProLeuIleSerIleValSerLeuAsnGlySerGlyThr 561
QY      1708 ATACAGGGCAACTCGGTGAAAACACAGAGTTCCTCCAAACGCTAGAGGCTGTAGAGCCA 1767
Db      562 ValGluGlyGlnLeuGlyThrAspAsnGluLeuIleMetSerArgAspLeuGluPro 581
QY      1768 TGCTGGCTPAATCATCGTATGATTTCTGTTGGATCCGGTATGCTTTTATGAAAC 1827
Db      582 CysValAlaAsnHisLysArgTyrPheLeuPheGlyHisThrValTyrValTyrGluAsp 601
QY      1828 TATAATTTTGTAAAGTGTAGACGCTGCCGATATACAGATTGCTAGCACATTTGTCGAG 1887
Db      602 TyrArgTyrValArgGluIleValHisAspValGlyMetIleSerThrTyrValAsp 621
QY      1888 CTTAATCTAACCTGCTAGAGATCGGAAATTTTCCTTTTCCCTTTTACACAAAGAA 1947
Db      622 LeuAsnLeuThrLeuLeuLysAspArgGluPheMetProLeuGlnValTyrThrArgAsp 641
QY      1948 GAGTTCGCGATGTTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2007
Db      642 GluLeuArgAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnMetHis 661
QY      2008 GAACCTAATTTATGACATAACAAGTAATAGAGTGGATGATGATGATGATGATGATGAT 2061
Db      662 SerLeuArgPheTyrAspIleAspLysValValGlnTyrAspSerGlyThrAlaIleMet 681
QY      2062 ---GGGCTGCAGGAATC----- 2076
Db      682 GlnGlyMetAlaGlnPheGlnGlyLeuGlyThrAlaGlyGlnAlaValGlyHisVal 701
QY      2077 -----GGCTGTATG-----TCATTACTAAAAAGATGCA 2106
Db      702 ValLeuGlyAlaThrGlyAlaLeuLeuSerThrValHisGlyPheThrThrPheLeuSer 721
QY      2107 AACCCAAATTAATGGCAACCAATTAAGAGCAGCGCAATGAGTTACAGATCTAATC 2166
Db      722 AsnProPheGlyAla-----LeuAlaValGlyLeuLeuValLeuAlaGlyLeuVal 738
QY      2167 AATGCTP-----AAAGCGATGACATTAGCTTACATACAGACTATGCCAAGATTGAAGCT 2220
Db      739 AlaAlaPhePheAlaTyrArgTyrValLeuLysLeu-LysThrSerPro---MetLysAl 757
QY      2221 AGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTTAAC-AATACTCTTAATGCAACATT 2279
Db      757 aLeuTyrProLeuThrThrLysGlyLeuLysGlnLeuProGluGlyMetAspProPheAl 777
QY      2280 AGAACAACTAAATGGCTAAACTAATTTAGATCAGCCATCAACCAAGCTAATACGA 2339
Db      777 aGluLysProAsnAlaThrAspThrProIleGluIleGlyAspSerGlnAsnThrGln 797
QY      2340 TAAACGACCTTTTGATTAATGAACACCAAAAT---TTAGTTGAGGACATACAAAGCACTAA 2396
;      :      :      :      :      :      :      :      :      :      :

```

```

Db      797 uProSerValAsnSerSerGlyPheAspProAspLysPheArgGluAlaGlnGluMetIleLy 817
QY      2397 A---ACCACCTTTAGAACAAAGTGTACTAACCTTGAAGCTTTGTGCATCAACTGCTTATAA 2453
Db      817 stYrMetThrLeuValSerAlaAlaGluArgGlnGlu-----Se 830
QY      2454 TCAAATTCGCAATTAATTTAGTGAATCATATCAATAAAGCTAGTAGTTTAACTAACTAAAC 2513
Db      830 rLysAlaArgLysLys-----AsnLysThrSerAlaLeuThrSerAr 845
QY      2514 ACTAGATCCACTA 2526
Db      845 gLeuThrGlyLeu 849

RESULT 10
US-09-338-326-21
; Sequence 21, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-326-21

Alignment Scores:
Pred. No.: 5,31e-151 Length: 868
Score: 1777.50 Matches: 375
Percent Similarity: 61.30% Conservative: 143
Best Local Similarity: 44.38% Mismatches: 259
Query Match: 31.38% Indels: 69
DB: 17 Gaps: 16

US-09-147-052-3 (1-3261) x US-09-338-326-21 (1-868)
QY 91 AGAGAGTTGTTTCGAGCGCTCCAGTTGCTGAGGAAGAGCTACGTTTATCTTTGTCGCC 150
Db 41 ArgGluAlaIleHisLysSerGlnAspAlaGluThrLysProThrPheTyrValCysPro 60

```



```

Db      757 aLeuTyrProLeuThrThrLysGlyLeuLysGlnLeuProGluGlyMetAspPropheAl 777
      ::::lll lll llllllllll ::::
Qy      280 AGAACATAAAATGGCTAAACTAATTTAGATCATGCCATCAACCAAGCTAATACGGA 2339
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      777 aGluLysProAsnAlaThrAspThrProIleGluGluIleGlyAspSerGlnAsnThrGl 797
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2340 TAAACAGACTTTTCATATGATGACACCCAAAT---TTAGTTGAACGATACAAAGCACATAA 2396
      :::: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      797 uProSerValasnSerGlyPheAspProAspLysPheArgGluAlaGlnGluMetIleLy 817
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2397 A---ACCATTAGAACCAACGCTACTTAACCTTGAAGGTTTGCATCACTGCTTATAA 2453
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      817 sTyrMetThrLeuValSerAlaAlaGluArgGlnGlu-----Se 830
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2454 TCAAAATCGCAATAATTTAGTGGATCATACAAATAAGCTAGTAGTTTAAATAACATAAAC 2513
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      830 rLysAlaArgLysLys-----AsnLysThrSerAlaLeuLeuThrSerAr 845
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2514 ACTAGATCCACTA 2526
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      845 gLeuThrGlyLeu 849

RESULT 11
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified Ttl-1 portion (downstream of BgII) of
; OTHER INFORMATION: pNZ40K-S
US-10-131-591A-23

Alignment Scores:
Pred. No.: 4,15e-151 Length: 357
Score: 1777.00 Matches: 352
Percent Similarity: 98.60% Conservative: 0
Best Local Similarity: 98.60% Mismatches: 5
Query Match: 31.37% Indels: 0
DB: 27 Gaps: 0

US-09-147-052-3 (1-3261) x US-10-131-591A-23 (1-357)

Qy      2155 ACAGATCTAACTAATGCTAAAGCGATGACATAGCTTCACTACAAAGACTATGCCAAGATT 2214
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      1 ThrAspLeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIle 20
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2215 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGCTTAACTAATACCTTAAATGCA 2274
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      21 GluAlaSerLeuSerSerAlaTyrSerGluAlaGluThrValAsnAsnLeuGlnAla 40
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2275 ACATTAGAACAACTAAATAATGGCTAAACCTAATTTAGATCAGCCATCAACCAAGCTAAT 2334
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      41 ThrLeuGluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsn 60
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2335 ACGGATAAAGCACTTTTGTATTAATGAACCCCAATTTAGTTGAAGCATACAAAGCACTA 2394
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      61 ThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeu 80
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2395 AAAACCACTTTAGAACACGCTGCTACTAACTTGAAGGTTTGTGCATCACTGCTTATAAT 2454
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      81 LysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerThrAlaTyrAsn 100
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::

```

```

Qy      2455 CAAATTCGCAATAATTTAGTGGATCTATACAAATAAGCTAGTAGTTAATAACTAAACA 2514
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      101 GlnIleArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThr 120
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2515 CTAGATCCACTAAATGGGGAAACGCTTTTACATCTCTAATGAGATTACTACAGCTAATAAG 2574
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      121 LeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLys 140
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2575 AATATTAAATATAGCTTATCAACTATTAAATGAACAAAAGACTAATGCTGATGCTATATCT 2634
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      141 AsnIleGlnAsnThrLeuSerThrIleAsnGlnGlnLysThrAsnAlaAspAlaLeuSer 160
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2635 AATAGTTTTTATAAAAAAGTCATCAAAATAATGAACAAAAGTTTGTAGGACACTTTTACA 2694
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      161 AsnSerPheIleLysLysValIleGlnAsnAsnGlnGlnSerPheValGlyThrPheThr 180
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2695 AACGCTAATGTTCAACCTTCAAACTACAGTTTGTCTTTTGTCTTTAGTCGTAGTACACACC 2754
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      181 AsnAlaAsnValGlnProSerGlnTyrSerPheValAlaPheSerAlaAspValThrPro 200
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2755 GTCAATTATAAATATGCAAGGACCGTTTGGATGCTGATGAACCTTCAAGTAGAATT 2814
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      201 ValAsnTyrLysTyrAlaArgThrValTrpAsnGlyAspGluProSerSerArgIle 220
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2815 CTTGCCAAACACAGATAGTATCACAGATGTTTCTGGATTTATAGTTTACGTGGACAAAC 2874
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      221 LeuAlaAsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsn 240
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2875 ACGAAGTACCAATTTAGTTTAGCAACTATGCTCCATCACTGGTGTATTTATATTCCT 2934
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      241 ThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPhePro 260
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2935 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTGATACAAATTAATAAT 2994
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      261 TyrLysLeuValLysAlaAlaAspAlaAsnValGlyLeuGlnTyrLysLeuAsnAsn 280
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      2995 GGAATGTTCAACAAGTTGAGTTTCCACTTCCACTAGTGCAGAAATAATACACTACACT 3054
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      281 GlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaGlnGlnThrThrAlaAsn 300
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      3055 CCAACTCCAGCAGTTGATGAGATTAAAGTTCTCTAAATCGTTTATCAGGTTTAAAGATT 3114
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      301 ProThrProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPhe 320
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      3115 GCCCAAAACACAACTCGAATTAAAGTTTCCAAACGGGTGAAGGAAATATGAATAAGTTGCG 3174
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      321 GlyGlnAsnThrIleGluLeuSerValProThrGlyGlyGlyAsnMetAsnLysValAla 340
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Qy      3175 CCAATGATTGGCAACATTTATCTTAGCTCAATGAAATAATGCTGATAAG 3225
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::
Db      341 ProMetIleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLys 357
      :lll:: lll :lll:: lll :lll:: lll :lll:: lll :lll::

RESULT 12
US-07-805-524-2
; Sequence 2, Application US/07805524
; GENERAL INFORMATION:
; APPLICANT: Babluk, Lorne
; APPLICANT: Van Der Hurk, Sylvia
; APPLICANT: Zamb, Tim
; APPLICANT: Fitzpatrick, David
; TITLE OF INVENTION: Bovine Herpesvirus Type 1 Polypeptides
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/805,524
FILING DATE: 19911211

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Gracey, Nancy J.
REGISTRATION NUMBER: 28,216

REFERENCE/DOCKET NUMBER: 293102000520
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792

TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 933 amino acids

TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
US-07-805-524-2

Alignment Scores:
Pred. No.: 2,08e-148 Length: 933
Score: 1749.00 Matches: 342
Percent Similarity: 65.268 Conservative: 105
Best Local Similarity: 49.938 Mismatches: 218
Query Match: 30.87% Indels: 20
DB: 3 Gaps: 6

US-09-147-052-3 (1-3261) x US-07-805-524-2 (1-933)

```

QY 52 GGTACGAACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAGTTGTTTCGAGCGTC 111
Db 97 GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116
QY 112 CAGTTGTCTGAGGAGAGTACGTTTATCTTTTCTTTGTCCTCCACCGAGTGGTTCAACCGTG 171
Db 117 GlnAlaAlaGlyGluAsnSerArgPheValCysProProSerGlyAlaThrVal 136
QY 172 ATCCGCTAGACCGCGCGGAAATGTCCCGAACCTAGAAAGCCACCGAGTGGGGTGA 231
Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
QY 232 GGAATGCGCATATTTTAAAGAGATATCATCTCCATATAATTTAAAGTGCAGCTTTAT 291
Db 157 GlyIleGlyValIleTyrGlyAsnIleAlaProTyrThrPheLysAlaTyrIleTyr 176
QY 292 TATAAAATATATTCAGACGACATGGACGGGACGACATATAGACAGATCACTAAT 351
Db 177 TyrLysAsnValIleValThrThrThrTrpAlaGlySerThrTyrAlaAlaIleThrAsn 196
QY 352 CGATATACAGTAGGACGCCGTTTCCATGTAAGAGATCAGCGATCTAATCGACGCCAA 411
Db 197 GlnTyrThrAspArgValProValGlyMetGlyIleThrAspLeuValAspLysLys 216
QY 412 GGAAGTGTCTATCAACAGACATACCTTAGAAACAATGTATATGTTGAACGGTTGAC 471
Db 217 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValValAlaPheAsp 236
QY 472 AGGGATGCGGGAGAAACAACTACTTCTTAAACCATCAAAATTCACACGCCCGCAATCT 531
Db 237 ArgAspAspProTrpGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 256
QY 532 AGGGATGCGCACACGACTAATGAGACGATATACCGTGTGGGGATACCATGATATATCGA 591
Db 257 ArgGlyTrpHisThrAspAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 276
QY 592 ACGGAACTCCGTAATGTATAGAGGAATGGATGGCGCTCTGTTTCGGTAT 651
Db 277 ThrGlyThrSerValAsnCysIleValGluValGluAlaArgSerValTyrProTyr 296
QY 652 TCATATTTTGAATGGCAATGGCGACATCGCGAACAATATCTCCATTTTATGGTCTATCC 711

```

```

Db 297 AspSerPheAlaLeuSerThrGlyAspIleIleTyrMetSerProPheTyrGlyLeuArg 316
QY 712 CCACGAGAGCTGCGCGACACCCATGGATATCCCGAGGATAATTTCAAAACATAGAT 771
Db 317 ---GluGlyAlaHisArgGluHisThrSerTyrSerProGluArgPheGlnIleGlu 335
QY 772 AGCTATTTTCAATGATTTGGACAGCGCTCGAAAGAACGCTTCCAGTCAAGCGTAAC 831
Db 336 GlyTyrTyrLysArgAspMetAlaThrGlyArgGluLeuLysGluProValSerArgAsn 355
QY 832 TTTTCTCATCATCATCATCTTCACAGTTGGTGGGACTGGCTCCCAAAACTACTCGTGA 891
Db 356 PheLeuArgThrGlnHisValThrValAlaIlePaspTirpValProLysArgLysAsnVal 375
QY 892 TGTTCATTAAGTAAAGAGGTGACTGAAATGTTGCGTCAACAGCTTAATGAGGACA 951
Db 376 CysSerLeuAlaLysTrpArgGluAlaAspGluMetLeuArgAspLysSerArgGlyAsn 395
QY 952 TACAGATTTATGCGCGTGAATTTGCGCAAGCTTTATCAGTAATACGACTAGTTGAT 1011
Db 396 PheArgPheThrAlaArgSerLeuSerAlaThrPheValSerAspSerHisThrPheAla 415
QY 1012 CCAATCGCATCATATTAGCAATGTTATTAACGCGGACGAGACACCAATCGAGCAG 1071
Db 416 LeuGlnAsnValProLeuSerAspCysValIleGluGluAlaGluAlaValGluArg 435
QY 1072 ATATTTAGGACAAATATATATGACAGTACGTCAGGTTGGACATGTACAATATTCTTG 1131
Db 436 ValTyrArgGluArgTyrAsnGlyThrHisValLeuSerGlySerLeuGluThrTyrLeu 455
QY 1132 GCTTCGCGGGGATTTATGTACATATCATGCTGTTCTATCAAAATCCCTGGCTCATATG 1191
Db 456 AlaArgGlyGlyPheValValAlaPheArgProMetLeuSerAsnGluLeuAlaLysLeu 475
QY 1192 TACCTCAGAGAATTGATGAGACAGACAGCGATGAGATGCTCCACCTGGTAAACAAT 1251
Db 476 TyrLeuGlnGluLeuAlaArgSerAsnGlyThrLeuGluGlyLeuPheAlaAla 493
QY 1252 AAGCATGCAATTTATAAGAAATGCTACCTCATTTGTCACGATTCGCGGAGATATCGA 1311
Db 494 ---AlaAlaProLysProGlyPro-----ArgArgAlaArgArgAlaAlaPro 508
QY 1312 AATGACCAAAATAGAAAATAACATTA-----CACCACACCACA 1350
Db 509 SerAlaProGlyGlyProGlyAlaAlaAsnGlyProAlaGlyAspGlyAlaGlyGly 528
QY 1351 GCTATTAATCGACATCGTCTGTTCAATTCGCGATGCTCCAATTTCTTTATGATCATATA 1410
Db 529 ArgValThrThrValSerSerAlaGluPheAlaAlaLeuGlnPheThrTyrAspHisIle 548
QY 1411 CAACCCATATTAATGATATGTTTAGTAGATTGCGCACAGCTTGGTGGCAATTCGAGAAT 1470
Db 549 GlnAspHisValAsnThrMetPheSerArgLeuAlaThrSerTrpCysLeuLeuGlnAsn 568
QY 1471 ACAGAACTGTTTATGCGCACCAAGGATTAAGATTAATCTCAGCGCTACACGAGTGA 1530
Db 569 LysGluArgAlaLeuTrpAlaGluAlaAlaLysLeuAsnProSerAlaAlaAlaSerAla 588
QY 1531 ACATTAGGAAGGAGAGTGGCTGCAAGAGTTGGGGATGTCGCTGCTATCGAGCTGC 1590
Db 589 AlaLeuAspArgArgAlaAlaAlaAlaAlaArgMetLeuGlyAspAlaMetAlaValThrTyrCys 608
QY 1591 ACTGCTATAGATGCGGAATCCGTCACCTTTGCAAAATTTCTATCGAGTTATCATCATCCT 1650
Db 609 HisGluLeuGlyGluGlyArgValPheIleGluAsnSerMetArg-----AlaProGly 626
QY 1651 AATACATGTTATAGCGGACCATTTGTTCTATTTTTCATATGGAGAAACCAAGAAACATA 1710
Db 627 GlyValCysTyrSerArgProProValSerPheAlaPheGlyAsnGluSerGluProVal 646
QY 1711 CAGGAGCAACTCGGTGAAACCAACAGAGTTGCTTCCAACGCTAGAGGCTCTAGAGCCATCG 1770

```



```

Db      647 GluGlyGlnLeuGlyGluAspAsnGluLeuLeuProGlyArgGluLeuValGluProCys 666
QY      1771 TCGGCTAATCATCGTAGATATTTCTGTTGGATCGGGTATGCTTTATTTGAAACATAT 1830
Db      667 ThrAlaAsnHisLysArgTyrPheArgPheGlyAlaAspTyrValTyrTyrGluAsnTyr 686
QY      1831 AATTTTGTAAAGTGTAGACGCTGCCGATATACAGATTGCTAGCACATTTCTCGAGCTT 1890
Db      687 AlaTyrValArgValProLeuAlaGluLeuGluValIleSerThrPheValAspLeu 706
QY      1891 AATCTAACCTGCTAGAAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAAGAAGAG 1950
Db      707 AsnLeuThrValLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu 726
QY      1951 TTGCGTGATGCTGGTGATTTGATGATTCAGAAAGTAGCTCGCCGCAATCACTACATGAA 2010
Db      727 LeuAlaAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisGlu 746
QY      2011 CTAAATTTTATGACATAACAAGTAATAGAGTGGATACAAATACCGC----- 2061
Db      747 LeuArgPheTyrAspIleAspArgValValIysThrAspGlyAsnMetAlaIleMetArg 766
QY      2062 GGGCTGCAGGAATTC 2076
Db      767 GlyLeuAlaAsnPhe 771

```

RESULT 13

```

US-08-096-183D-4
; Sequence 4, Application US/08096183D
; GENERAL INFORMATION:
; APPLICANT: Roger K. Maes and Stephen J. Spatz
; TITLE OF INVENTION: Recombinant Foxvirus
; TITLE OF INVENTION: Vaccine Against
; TITLE OF INVENTION: Feline Rhinotracheitis
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,183D
; FILING DATE: July 26, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-166
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: None
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: Polypeptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO

```

```

; ORIGINAL SOURCE:
; ORGANISM: Feline herpesvirus-1
; STRAIN: 1
; INDIVIDUAL ISOLATE: C-27
; CELL TYPE: N/A
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: Deduced sequence
; OTHER INFORMATION: gb
; US-08-096-183D-4

```

```

Alignment Scores:
Pred. No.: 7, 28e-148 Length: 943
Score: 1743.00 Matches: 377
Percent Similarity: 57.54% Conservative: 161
Best Local Similarity: 40.32% Mismatches: 267
Query Match: 30.77% Indels: 130
DB: 4 Gaps: 21

```

```

US-09-147-052-3 (1-3261) x US-08-096-183D-4 (1-943)

```

```

QY      7 TATTTTAGCGGAATTGCATATTT----- 30
Db      22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
QY      31 -----TTCCTTATAGTT 42
Db      42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrrPile 61
QY      43 ATCTATAT-----GGTAGAACATCATCTCCG 69
Db      62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY      70 ACTACCCAAAATGTGCATCAAGAGAGTT----- 99
Db      82 ProArgArgThrValAlaThrProGluValGlyValHisHisGlnAsnGlnLeuGlnIle 101
QY      100 -----GTTTCGAGCGTCCAGTTGTCTGAGGAAGAG----- 129
Db      102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
QY      130 TCTAGCTTTATCTTTGTCCTCCACAGTGGGTTCACCGTGATCGGTCTAGAACCGCG 189
Db      122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
QY      190 CGAAAATGTCCGGAACCTAGAAAAGCCACCGAGTGGGTGAAGGAATCCGATATATTAT 249
Db      142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
QY      250 AAAGAGATATCATGCTCCATATAAATTTAAAGTCACGCTTTTATATATAAATATCATTCAG 309
Db      162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
QY      310 ACGACGACATGACGGGGAGACATATAGACAGATCACTAATCGATATACAGATGAGCG 369
Db      182 ThrThrValTyrPheSerGlySerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201
QY      370 CCCGTTTCCATTGAAGACATCAGGATCAATTCGACGGCAAGAGAGATGCTCATCTAA 429
Db      202 ProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeuSerLys 221
QY      430 GCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTTCAGAGGATCGGGAGAGAAA 489
Db      222 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241
QY      490 CAAGTACTTCTAAACCATCAAAATTCACACCGCCGCAATCTAGGCGATGGCAGCAGCT 549
Db      242 GluLeuProLeuLysProSerSerThrLeuSerArgValArgGlyTrrPheIle---Thr 260
QY      550 AATGAGACGTATACCGTGTGGGATCACCATCGATATATCGAACGGGAACCTCCGCTCAAT 609
Db      261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisHisSerGlyThrSerValAsn 280

```

QY	610	TGTATAGTAGAGAAATGGATGCCCGCTCTGTTTCGGTATTCATATTTTTCGAATGCC	666
DB	281	CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer	300
QY	670	AATGGCGACATCGCGACATATCTCCATTTTATGGTCTATCCCCCAGAGGCTCGCGCA	729
DB	301	ThrGlyAspValIleHisMetSerProPhePheGlyLeuArg---AspGlyAlaHisVal	319
QY	730	GAACCCATGGGATATCCCCAGGATAAATTCAAACAACACTAGTACTATTTTCAATGGAT	789
DB	320	GluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrTyrProIleAsp	339
QY	790	TTGACAAAGCGTCCGAAAAGCAAGCCTTCCAGTCAAGCGTAACCTTTCATCACAATCACAC	849
DB	340	LeuAspThrAspTyrThrGlyAla---ProValSerArgAsnPheLeuGluThrProHis	358
QY	850	TTACAGCTTGGGTGGGACTGGGTCGCAAAACACTACTCGTGTATGTTCAACTGACTAAGTG	909
DB	359	ValThrValAlaTrpAsnTrpThrProLysSerGlyArgValCysThrLeuAlaLysTrp	378
QY	910	AAAGAGTGACTGAATTTGGGTGCAACAGTAAATGGAGATACACATTTATGGCCCGT	969
DB	379	ArgGluIleAspGluMetLeuPrometAsnIle---GlySerTyrArgPheThrAlaLys	397
QY	970	GAACCTTTCGGCAAGCTTTATCAGTAATACGACTCAGTTTGCATCCAAATCGCATCATATTA	1029
DB	398	ThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeu	417
QY	1030	GGACAATGATTAACCCGGAGGCAAGACGAAATCGAGCAGATATTTTAGCAACAATAT	1089
DB	418	GlyAspCysAlaThrLysGluAlaAlaGluAlaIleAspArgIleTyrLysSerLysTyr	437
QY	1090	AATGACAGTCACGTCGAAGTTGGCAGTGTACATATTTCTTGCTCTCGGGGATTTATT	1149
DB	438	SerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeu	457
QY	1150	GTACATATCAGCCTGTTCTATCCAAATCCCTGGCTCATATGTACCTTCAGAGAATTCATG	1209
DB	458	IleAlaPheArgProMetIleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeuAla	477
QY	1210	AGACACAACAGGACCGATGAGATCTCGACCTGGTAAACAATAGCATGCAATTTATPAAG	1269
DB	478	ArgSerAsnArgThr-----ValValAspLeu-----SerAlaLeuLeuAsn	491
QY	1270	AAAAATGCTACCTTCATGTCACGATTCGCGCGAGATATTCGAATGCCACAATAGAAAA	1329
DB	492	ProSerGlyGluThrValGlnArgThrArgSerValProSerAsnGlnHisArg	511
QY	1330	ATA-----ACATTAGACACACACACGACTATTAAA	1359
DB	512	SerArgArgSerThrIleGluGlyGlyIleGluThrValAsnAsnAlaSerLeuLeuLys	531
QY	1360	TCGACATCGTCTGTCAATTCGCCGACGCTCCAAATTTCTTATGATCATATACAAACCCAT	1419
DB	532	ThrThrSerSerValGluPheAlaMetLeuGlnPheAlaTyrAspTyrIleGlnAlaHis	551
QY	1420	ATTAATGATATGTTAGTAGGATTCGCCAGCTTGGTGGCAATTCGAGAATAGAGAACTT	1479
DB	552	ValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysThrLeuGlnAsnArgGluHis	571
QY	1480	GTTTTATGGCAACGAGGATAAAGATTAACTTCAGCGCTACACGAGTGCAACATTAGGA	1539
DB	572	ValLeuTrpThrGluThrLeuLysLeuAsnProGlyGlyValValSerMetAlaLeuGlu	591
QY	1540	AGGAGAGTGGCTGCAAGATGTTGGGGGATGTCGCTGCTATCGAGCTGACTGCTATA	1599
DB	592	ArgArgValSerAlaArgLeuLeuGlyAspAlaValAlaValThrGlnCysValAsnIle	611
QY	1600	GATGCGGAATCCGTCACCTTGC AAAATTCATCGCGAGTTATCACATCCACTAATACATGT	1659
DB	612	SerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrThrCys	631

QY	1660	TATAGCCGACCATGGTGTCTTATTTTCATATGGAGAAAACCAAGAAACATACATACAGGGACAA	1719
Db	632	TyrSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluTyrIleGluGlyGln	651
QY	1720	CTCGGTGAAACCAACGAGTGTCTCCCAACCGCTAGAGCGCTAGAGCCATGCTCGGCTAAT	1779
Db	652	LeuGlyGluAsnAsnGluLeuLeuValGluArgLysLeuIleGluProCysThrValAsn	671
QY	1780	CATCGTAGATATATTTCTGTGTGGATCCGGTTATGCTTTATTTCAAAACATAATATTTGTT	1839
Db	672	AsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyrVal	691
QY	1840	AAGATGGTAGACCGTCGCCGATATACAGATTCGTAGCACATTTCTGAGCTTAATCTAAACC	1899
Db	692	ArgLysValProLeuSerGluIleGluLeuIleSerAlaTyrVal---IleLysSerThr	710
QY	1900	CTGCTAGAAGATCGGGAATTTTGGCTTTATCCGTTTACACAAAGAACGAGTTCGCTGAT	1959
Db	711	LeuLeuGluAspArgGluPheLeuHis---SerSerTyrThrArgAlaGluLeuGluAsp	729
QY	1960	GTTGGGTATTTGGATTATGCAGAAGTAGCTCGCGCAACCACTACATGAACCTAAATTT	2019
Db	730	ThrGlyProPheAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuLysPhe	749
QY	2020	TATGACATAAACAAGTAATAGAACTGGATACAAAT-----	2055
Db	750	TyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMetAla	769
QY	2056	-----TACGCGGGGTGCGAGGAATTCGGCTGT-----	2082
Db	770	AsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGlyAla	789
QY	2083	-----ATGCTATTACTAAAGATGCCAAACCCAAATAATATGCGCCAAACC	2127
Db	790	AlaSerAlaValIleSerThrValSerGlyValSerSerPheLeuAsnAsnPropGly	809
QY	2128	CAATTAGAAGCACCGCGCAATGGAGTTAACAGATCTAATCAATGCT-----	2172
Db	810	AlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleValAlaAlaPheLeuAlaTyrArg	829
QY	2173	-----AAAGCGATG-----ACATTAGCTTCA	2193
Db	830	TyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArgAsn	849
QY	2194	CTACAAGACTATGCCAAGATTAAGCTAGTTTATCATCTGCTTATAGTAGCGCTGAACA	2253
Db	850	LeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyLysPheSerAspProGlyVal	869
QY	2254	GTTTAAACAATAACCTTAATGCAACATTTAGACAACACTAAAATGGCTAAA---ACTAATTTA	2310
Db	870	AspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSerLeu	889
QY	2311	GAATCAGCATCAACCAAGCTAATACGGATAAAACGACTTTTGATTAATCAACCCCAAT	2370
Db	890	ValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyProAla	908
QY	2371	TTAGTTGAAGCATAC-----AAAGCATATAAAACC-----ACTTTAGAACAACT	2415
Db	909	IleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgArgGlyProLysTyrGlnArg	928
QY	2416	GCTACTAAACCTTGAAGCTTTGTCATCACTGCTTATAATCAAAATT	2460
Db	929	LeuAsnAsnLysSerGlyAspAspThrGluThrAsnLeuVal	943

RESULT 14

```

US-09-521-738-14
RESOLVED 14
; Sequence 14, Application US/09521738
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE

```

;
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9
; CURRENT APPLICATION NUMBER: US/09/521,738
; CURRENT FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
; US-09-521-738-14

Alignment Scores:

Prod. No.: 7,28e-148 Length: 943
Score: 1743.00 Matches: 377
Percent Similarity: 57.54% Conservative: 161
Best Local Similarity: 40.32% Mismatches: 267
Query Match: 30.77% Indels: 130
DB: 19 Gaps: 21

US-09-147-052-3 (1-3261) x US-09-521-738-14 (1-943)

```
QY 7 TATTTAGCGGAATTGCATATT----- 30
Db 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyLeuAlaAlaThrGlySerArg 41
QY 31 -----TTCCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCTATAT-----GGTACGACATCATCTCCG 69
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 70 AGTACCCAAAATGTGACATCAAGAGATT----- 99
Db 82 ProArgThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
QY 100 -----GTTTCGAGCGTCCAGTTCTCTGAGAAAGAG----- 129
Db 102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
QY 130 TCTPACGTTTATCTTGTCCCCACCAGTGGGTCAACCGTGATCGTCTAGAACCGCG 189
Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
QY 190 CGAAAATGTCCGAACCTAGAAAAGCCACCGAGTGGGTGAAGAAATCGCATATTATT 249
Db 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyLeuAlaValIlePhe 161
QY 250 AAGAGAATATCATGTCATATAATTTAAAGTGACGCTTTATTATATAAATATCATTCAG 309
Db 162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
QY 310 ACGACCATATGAGCGGAGCATATAGACAGATCACTAATCGATATACAGATAGACG 369
Db 182 ThrThrValTrpSerGlySerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201
QY 370 CCGTTTTCATTAAGAGATACGGATCTAATCGACGGCAAGAGAGTCTCATCTAA 429
Db 202 ProValLysValGlnGluIleThrAspLeuIleAspArgArgGlyMetCysLeuSerLys 221
QY 430 GCAAGATACCTTAGAAACAATGTATATGTTGAAGCGCTTTGACAGGGATCGGGAGAAAA 489
Db 222 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241
QY 490 CAAGTACTTCTAAACCAATAAATTCACACCGCCGAATCTAGGCCATGGCACGACT 549
Db 242 GluLeuProLeuLysProProSerThrLeuSerArgValArgGlyTrpHis---Thr 260
QY 550 AATGACACTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACCTCCGTCAT 609
Db 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisSerGlyThrSerValAsn 280
```

```
QY 610 TGTATAGTAGAGAAATGGATGCCCTCTGTGTTTCCGTATTTCATATTTCGAATGGCC 669
Db 281 CysIleValGluGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300
QY 670 AATGGCGACATCGGAACATATCTCCATTTTATGGTCTATCCCCACGAGGTCGCCGA 729
Db 301 ThrGlyAspValIleHisMetSerProPhePheGlyLeuArg---AspGlyAlaHisVal 319
QY 730 GAACCCATGGATATCCCGAGGATAATTTCAAAACAACATAGATAGCTATTTTCAATGGAT 789
Db 320 GluHisThrSerTyrSerSerArgPheGlnGlnIleGluGlyTyrProIleAsp 339
QY 790 TTGGACAACGCTCGAAAAGCAAGCCCTTCAGTCAACGGTAACCTTCTCATCATCACAC 849
Db 340 LeuAspThrAspTyrThrGlyAla---ProValSerArgAsnPheLeuGluThrProHis 358
QY 850 TTCACAGTTGGGTGGGACTGGGCTCCAAAACACTCTGTATGTTCATGTAAGTGG 909
Db 359 ValThrValAlaTrpAsnTrpThrProLysSerGlyArgValCysThrLeuAlaLysTrp 378
QY 910 AAAGAGTGACTGAAATGTTGCTGCAACAGATTAAATGGGAGATACAGATTTATGGCCGT 969
Db 379 ArgGluIleAspGluMetLeuProMetAsnIle---GlySerTyrArgPheThrAlaLys 397
QY 970 GAACCTTTCGCAACGTTTATCAGTAATACGACTGAGTTGATCCAAATCGCATCATATTA 1029
Db 398 ThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeu 417
QY 1030 GGACATGTATTAACCGCAGGAGCAAGCAATCGACAGATATAGGACAAAATAT 1089
Db 418 GlyAspCysAlaThrLysGluAlaIleGluAlaIleAspArgIleTyrLysSerLysTrp 437
QY 1090 AATGACAGTCACGTCGAGTGGACATGTACAAATTTCTTGGCTCTCGGGGATTTAT 1149
Db 438 SerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeu 457
QY 1150 GTAGCATATCAGCCTGTTCTTCCAAATCCCTGGCTCATATGTTACCTCAGAGAAATGATG 1209
Db 458 IleAlaPheArgProMetIleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeuAla 477
QY 1210 AGACAACAACGACCCGATGAGATGCTGACCTGGTAAACAATAAGCATCAATTTATAAG 1269
Db 478 ArgSerAsnArgThr-----ValValAspLeu-----SerAlaLeuLeuAsn 491
QY 1270 AAAAATGCTACCTCATTTGTCACGATTCGCGCGAGATATTGGAATGCAACCAATAGAAA 1329
Db 492 ProSerGlyGluThrValGlnArgThrArgArgSerValProSerAsnGlnHisArg 511
QY 1330 ATA-----ACATTTAGACACACACACGCTATTATAA 1359
Db 512 SerArgArgSerThrIleGluGlyGlyIleGluThrValAsnAsnAlaSerLeuLeuLys 531
QY 1360 TCGACATCGTCTCTCAATTCGCATCTCCCAATTTCTTATCATCATATACAAACCCAT 1419
Db 532 ThrThrSerSerValGluPheAlaMetLeuGlnPheAlaTrpCysThrLeuGlnAsnArgGluHis 551
QY 1420 ATTAATGATATGTTTATGAGATTGCCACAGCTTGGTGGCAATTCGAGATAGAGAACTT 1479
Db 552 ValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysThrLeuGlnAsnArgGluHis 571
QY 1480 GTTTTATGCAACGAGGATAAAGATTAATCTTCAGCGCTACACGCGCTACACAGTGCACATAGGA 1539
Db 572 ValLeuTrpThrGluThrLeuLysLeuAsnProGlyGlyValValSerMetAlaLeuGlu 591
QY 1540 AGGAGAGTGGCTGCAAAAGATGTTGGGGATGTCGCTGCTATCGAGCTGCACATAGTA 1599
Db 592 ArgArgValSerAlaArgLeuLeuGlyAspAlaValAlaValThrGlnCysValAsnIle 611
QY 1600 GATCGGGAATCCGTCACCTTTCGCAAAATCTATCGAGTGTATCATCTCCACTAATCATGT 1659
Db 612 SerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrThrCys 631
```

```
QY 1660 TATAGCCGACCATTTGCTTCTATTTTCATATGAGAAAAACAGGAAACATACAGGACAA 1719
Db 1661 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 TyrSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluTyrIleGluGlyGln 651
QY 1720 CTCGGTGAACAACAGAGTCTGCTTCCACAGCTAGAGGCTGTAGAGCCATGCTCGCTAAT 1779
Db 1721 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 LeuGlyGluAsnAsnGluLeuValGluArgLysLeuIleGluProCysThrValAsn 671
QY 1780 CATCGTAGATATTTCTGTTGGATCGGTTATGCTTTATTTGAAACTATAATTTGTT 1839
Db 1781 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
672 AsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyrVal 691
QY 1840 AAGATGGAGACGCTGCGCATACAGATTGCTGACATTTGCGAGCTTAATCTAAC 1899
Db 1841 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
692 ArgLysValProLeuSerGluIleGluLeuSerAlaTyrVal---IleLysSerThr 710
QY 1900 CTGCTAGAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAGAGAGTTCGGTGAT 1959
Db 1901 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
711 LeuLeuGluAspArgGluPheLeuHis---SerSerTyrThrArgAlaGluLeuGluAsp 729
QY 1960 GTTGGTGATTGGATTATGCAAGATAGCTGCGCGCATCACTACATGAACTTAATTT 2019
Db 1961 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
730 ThrGlyProPheAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuLysPhe 749
QY 2020 TATGACATAAACAAGTAATAAGATGGATACAAAT----- 2055
Db 2021 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
750 TyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMetAla 769
QY 2056 -----TACGGCGGGCTGCAGGAATTCGGCTGT----- 2082
Db 2057 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
770 AsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGlyAla 789
QY 2083 -----ATGCTATTACTAATAAAGATCAACACCAATATATGCGCAAC 2127
Db 2084 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
790 AlaSerAlaValIleSerThrValSerGlyValSerPheLeuAsnAsnProPheGly 809
QY 2128 CAATTAGAAGCAGCGCAAGTGAAGTAAACAGATCAATCAATGCT----- 2172
Db 2129 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
810 AlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleValAlaAlaPheLeuAlaTyrArg 829
QY 2173 -----AAACCGATG-----ACATTAGCTTCA 2193
Db 2174 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
830 TyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArgAsn 849
QY 2194 CTACAAGACTATGCAAGATGAAGTAGTTTATCATCTGCTTATAGTGAAGCTGAACA 2253
Db 2195 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
850 LeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyLysAspSerAspProGlyVal 869
QY 2254 GTTAACAATAACCTTAATGCAACATTAACAACATAAATAATGGCTAAA---ACTAATTTA 2310
Db 2255 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
870 AspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSerLeu 889
QY 2311 GAATCAGCCATCAACCAAGCTAATACGGATAAAGCACTTTTGTATATGAACACCAAT 2370
Db 2312 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
890 ValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyProAla 908
QY 2371 TTAGTTGAAGCATAC-----AAAGCACTAAAAACC-----ACTTTAGAACAACGT 2415
Db 2372 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
909 IleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgGlyProLysTyrGlnArg 928
QY 2416 GCTACTAACCTTGAAGCTTTGTCATCACTCACTTATAATCAAAAT 2460
Db 2417 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
929 LeuAsnAsnLeuAspSerGlyAspAspThrGluThrAsnLeuVal 943
```

RESULT 15

```
US-08-541-878-6
; Sequence 6, Application US/08541878
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Black, Darla
; APPLICANT: Scnicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
```

```
; TITLE OF INVENTION: Virus Genes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cox & Smith Incorporated
; STREET: 112 East Pecan Street, Suite 2000
; CITY: San Antonio
; STATE: Texas
; COUNTRY: USA
; ZIP: 78205
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,878
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/042,747
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haymond, W. Bradley
; REGISTRATION NUMBER: 35186
; REFERENCE/DOCKET NUMBER: S-0072.179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 210-554-5500
; TELEFAX: 210-226-8395
; TELEX: 767609
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 891 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-541-878-6
Alignment Scores:
Pred. No.: 1.48e-147 Length: 891
Score: 1739.50 Matches: 332
Percent Similarity: 64.16% Conservative: 103
Best Local Similarity: 48.97% Mismatches: 216
Query Match: 30.71% Indels: 27
DB: Gaps: 3
```

US-09-147-052-3 (1-3261) x US-08-541-878-6 (1-891)

```
QY 52 GGTACGAACATCATCTCCGAGTACCCAAAATGTGACATCAAGAGAAGTTGTTTCGAGGTC 111
Db 53 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72 GlyThrAsnAlaSerValGluAlaGlyHisAlaThrLeuArgGluAsnLeuArgAspIle 91
QY 112 CAGTTGCTCGAGGAAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTCAACCGTG 171
Db 113 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
92 LysAlaLeuAspGlyAspAlaThrPheTyrValCysProProThrGlyAlaThrVal 111
QY 172 ATCCGCTACAGCCGCGCGGCAAAATGTCCTCGAACCTAGAAAGACCCAGTGGGGTGAA 231
Db 173 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 ValGlnPheGlnGlnProArgProCysProArgAlaProHisGlyGlnAsnTyrThrGlu 131
QY 232 GGAATCGCGATATATTAAAGAGAATATCAGTCCATATAAATTTAAAGTGACGCTTAT 291
Db 233 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 GlyIleAlaValIlePheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyr 151
QY 292 TATAAATAATCATTCAGACGACATGGACGGGACGACATATAGACAGATCACTAAT 351
Db 293 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
152 TyrLysAspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGly 171
QY 352 CGATACACAGATAGACGCGCGTTTCATTTGAGAGATCAAGATCACTAATCGACGCCAA 411
Db 353 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
172 IlePheGluAspArgAlaProValProPheGluGluValIleAsnLysIleAsnAlaArg 191
QY 412 GGAAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGGTTTGAC 471
Db 413 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 192 GlyValCysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHis 211
QY 472 AGGATGCGGAGAGAAAACAAGTACTTCTAAACATCAAAATCAACACGCCGGAATCT 531
Db 212 ArgAspAspGluSerAspMetLysLeuLysProAlaLysAlaAlaThrArgThrSer 231
QY 532 AGGCATGGCAGACACTAAATGAGACGATACCGTGTGGGATCACCATGATATATCGA 591
Db 232 ArgGlyTrpHisThrAspLeuLysTyrAsnProSerArgIleGluAlaPheHisArg 251
QY 592 ACGGAACCTCGTCAATTTATAGTAGAGAAATGGATGCCGCTCTGTGTTCCGCTAT 651
Db 252 TyrGlyThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 271
QY 652 TCATATTTTGCATGGCCCAATGGGACATCGCGAACATATCTCCATTTTATGGTCTATCC 711
Db 272 AspGluPheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg 291
QY 712 CCACGAGAGCTGCGCAGAACCCATGGGATATCCCGAGGATATTTTCAAACTAGAT 771
Db 292 --- AspGlyAlaHisAlaGluHisThrAlaTyrAlaAlaAspArgPheArgGlnValAsp 310
QY 772 AGCTATTTTCAATGATTTGGACAAGCGTCGAAAGCAAGCCTCCAGTCAAGCGTAAC 831
Db 311 GlyTyrTyrGluArgAspLeuSerThrGlyArgArgAlaSerThrProAlaThrArgAsn 330
QY 832 TTTCTCATCATCACATTCACAGTTGGGTGGGACTGGGCTCCAAAACACTACTCGTGA 891
Db 331 LeuLeuThrThrProLysPheThrValGlyTrpAspTrpAlaProLysArgProSerVal 350
QY 892 TGTCAATGACTTAAGTGAAGAGGTGACTCAATATGTTGCGTGCACACAGTTAATGGGAGA 951
Db 351 CysThrLeuThrLysTrpGlnGluValAspGluMetLeuArgAlaGluTyrGlyProSer 370
QY 952 TACAGATTTATGCGCCGTGAACCTTTCGGCAACGTTTATCATCAGTACGACTGAGTTGAT 1011
Db 371 PheArgPheSerSerAlaLeuSerThrThrPheThrThrAsnArgThrGluTyrAla 390
QY 1012 CCAATCCCATCATATTAGGACATGTATTAAACCGCAGGAGAGAGCAATCAGCAGAG 1071
Db 391 LeuSerArgValAspLeuGlyAspCysValGlyArgGluAlaArgGluAlaValAspArg 410
QY 1072 ATATTAGGACAAAATAATGACAGTCACGTCAAGTTGGACATGTACATATTTCTTG 1131
Db 411 IlePheLeuArgTyrAsnGlyThrHisValLysValGlyGlnValGlnTyrTyrLeu 430
QY 1132 GCTCTCGGGGATTTATAGCATATCAGCCTGTCTTATCCAAATCCCTCGCTCATATG 1191
Db 431 AlaThrAlaGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnGlyLeuValGluLeu 450
QY 1192 TACCTCAGAGATTTGATGAGACACACAGGCCGATGAGATGCTGCACCTGGTAACAAT 1251
Db 451 TyrValArgGluLeuLeuArgGluGlnGluArgPro----- 463
QY 1252 AAGCATGCAATTATAGAAAATGCTACCTCATTTGTCACGATTCGCGGAGATATTCGA 1311
Db 464 -----GlyAspAlaAla 467
QY 1312 AATGCACCAATAGAAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCT 1371
Db 468 AlaThrProLysProSerAlaAspProProAspValGluArgIleLysThrThrSerSer 487
QY 1372 GTTCAATTCGCATGCTCCATTTCTTATGATCATATACAAACCCATATTAAATGATATG 1431
Db 488 ValGluPheAlaArgLeuGlnPheThrTyrAspHisIleGlnArgHisValAsnAspMet 507
QY 1432 TTTAGTAGATTCACACAGCTTGTGCGCAATTCAGAGATAGAGACTGTTTATGGCAC 1491
Db 508 LeuGlyArgIleAlaIleAlaTrpCysGluLeuGlnAsnHisGluLeuThrLeuTrpAsn 527
QY 1492 GAAGGGATAAAGATTAATCTTAGCGCTACAGCGAGTCAACATTAGGAAGGAGAGTGGCT 1551
Db 528 GluAlaArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgValSer 547

QY 1552 GCAAGATGTTGGGGATGTCGCTGCTCTATCGAGCTGCACCTGCTATAGATCGCGAATCC 1611
Db 548 AlaArgMetLeuGlyAspValMetAlaValSerThrCysValProValThrProAspAsn 567
QY 1612 GTCACCTTTCGAAAATTCATCGGAGTATCATCCACTCACTAACTACATGTTATAGCCGACA 1671
Db 568 ValIleMetGlnAsnSerMetArgValProAlaArgProGlyThrCysTyrSerArgPro 587
QY 1672 TTGGTTTCATTTTCATATGGAGAAACCAAGGAAACATACAGGGACAACTCGTGGTAAAC 1731
Db 588 LeuValSerPheArgTyrGluGluGlyGlyProLeuValGluGlyGlnLeuGlyGluAsp 607
QY 1732 AACGATGCTTCCAACGCTAGAGCTGTAGAGCATGCTCGGCTAACTCATCGTAGATAT 1791
Db 608 AsnGluIleArgLeuGluArgAspAlaLeuGluProCysThrValGlyHisArgArgTyr 627
QY 1792 TTTCTGTTTGGATCCGGTTATGCTTTATTGAAAACATAAATTTGTTAAGATGCTAGAC 1851
Db 628 PheThrPheGlyAlaGlyTyrValTyrPheGluAspTyrAlaTyrSerHisGlnLeuGly 647
QY 1852 GCTGCCGATATACAGATTGTCAGCACATTTGTCGAGCTTAATCTAACCTGCTAGAAGAT 1911
Db 648 ArgAlaAspValThrThrValSerThrPheIleAsnLeuAsnLeuThrMetLeuGluAsp 667
QY 1912 CGGGAATTTGCTTTTATCCGTTTACACAAAGAGAGTTCGCTGATGTTGCTGTATG 1971
Db 668 HisGluPheValProLeuGluValTyrThrArgGlnGluIleLysAspSerGlyLeuLeu 687
QY 1972 GATTATGCAAGTAGCTCGCGCAATCAACTACATGAACCTTAAATTTTATCACAATAC 2031
Db 688 AspTyrThrGluValGlnArgAsnGlnLeuHisAlaLeuArgPheAlaAspIleAsp 707
QY 2032 AAAGTAATAGAGTGGATACAAAT-----TACGCGGGCTGCAGGAATTC 2076
Db 708 ThrValIleLysAlaAspAlaHisAlaProLeuPheAlaGlyLeuTyrSerPhe 725

Search completed: October 8, 2003, 19:15:21
Job time : 1059.75 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 18:12:37 ; Search time 178.635 Seconds
(without alignments)
5882.831 Million cell updates/sec

Title: US-09-147-052-3

Perfect score: 5665

Sequence: 1 atgcactattttaggcgga.....gacccgggtacattttataa 3261

Scoring table: BLOSUM62

-Q- /cgn2_1/USPTO_spool/US09147052/runat_08102003_154344_29960/app.query.fasta_1.5980
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 1201306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=alp
-Q-/cgn2_1/USPTO_spool/US09147052/runat_08102003_154344_29960/app.query.fasta_1.5980
-DB=Published Applications_AA -OFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPECI=0 -LOOPECI=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09147052.ecgn_1_1_355 -Runat_08102003_154344_29960
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	5619	99.2	1086	9	US-09-147-052-4	Sequence 4, Appl1
2	2015	35.6	456	9 <th>US-09-147-052-2</th> <th>Sequence 2, Appl1</th>	US-09-147-052-2	Sequence 2, Appl1
3	2008	35.4	456	12 <th>US-09-901-572A-3</th> <th>Sequence 3, Appl1</th>	US-09-901-572A-3	Sequence 3, Appl1
4	1929.5	34.1	384	15 <th>US-10-131-591A-12</th> <th>Sequence 12, Appl1</th>	US-10-131-591A-12	Sequence 12, Appl1
5	1777	31.4	357	15 <th>US-10-131-591A-23</th> <th>Sequence 23, Appl1</th>	US-10-131-591A-23	Sequence 23, Appl1
6	1736	30.6	904	9 <th>US-09-894-998-18</th> <th>Sequence 18, Appl1</th>	US-09-894-998-18	Sequence 18, Appl1
7	1736	30.6	904	12 <th>US-10-200-562-18</th> <th>Sequence 18, Appl1</th>	US-10-200-562-18	Sequence 18, Appl1
8	1736	30.6	904	12 <th>US-10-237-551-18</th> <th>Sequence 18, Appl1</th>	US-10-237-551-18	Sequence 18, Appl1
9	1736	30.6	904	15 <th>US-10-121-988-18</th> <th>Sequence 18, Appl1</th>	US-10-121-988-18	Sequence 18, Appl1
10	778.5	13.7	907	15 <th>US-10-223-538-2</th> <th>Sequence 2, Appl1</th>	US-10-223-538-2	Sequence 2, Appl1
11	758.5	13.4	830	14 <th>US-10-055-364-45</th> <th>Sequence 45, Appl1</th>	US-10-055-364-45	Sequence 45, Appl1
12	754	13.3	865	14 <th>US-10-055-364-24</th> <th>Sequence 24, Appl1</th>	US-10-055-364-24	Sequence 24, Appl1
13	750	13.2	793	14 <th>US-10-055-364-55</th> <th>Sequence 55, Appl1</th>	US-10-055-364-55	Sequence 55, Appl1
14	738.5	13.0	844	14 <th>US-10-055-364-40</th> <th>Sequence 40, Appl1</th>	US-10-055-364-40	Sequence 40, Appl1
15	725.5	12.8	808	14 <th>US-10-055-364-38</th> <th>Sequence 38, Appl1</th>	US-10-055-364-38	Sequence 38, Appl1
16	719	12.7	359	14 <th>US-10-055-364-48</th> <th>Sequence 48, Appl1</th>	US-10-055-364-48	Sequence 48, Appl1
17	716	12.6	823	14 <th>US-10-055-364-37</th> <th>Sequence 37, Appl1</th>	US-10-055-364-37	Sequence 37, Appl1
18	709.5	12.5	829	14 <th>US-10-055-364-46</th> <th>Sequence 46, Appl1</th>	US-10-055-364-46	Sequence 46, Appl1
19	704	12.4	831	14 <th>US-10-055-364-39</th> <th>Sequence 39, Appl1</th>	US-10-055-364-39	Sequence 39, Appl1
20	700.5	12.4	818	14 <th>US-10-055-364-44</th> <th>Sequence 44, Appl1</th>	US-10-055-364-44	Sequence 44, Appl1
21	695	12.3	791	14 <th>US-10-055-364-42</th> <th>Sequence 42, Appl1</th>	US-10-055-364-42	Sequence 42, Appl1
22	695	12.3	792	14 <th>US-10-055-364-41</th> <th>Sequence 41, Appl1</th>	US-10-055-364-41	Sequence 41, Appl1
23	680	12.0	824	14 <th>US-10-055-364-43</th> <th>Sequence 43, Appl1</th>	US-10-055-364-43	Sequence 43, Appl1
24	595.5	10.5	206	14 <th>US-10-127-733-2</th> <th>Sequence 2, Appl1</th>	US-10-127-733-2	Sequence 2, Appl1
25	343.5	6.1	195	14 <th>US-10-055-364-2</th> <th>Sequence 2, Appl1</th>	US-10-055-364-2	Sequence 2, Appl1
26	314	5.5	62	15 <th>US-10-131-591A-6</th> <th>Sequence 5, Appl1</th>	US-10-131-591A-6	Sequence 5, Appl1
27	302	5.3	62	15 <th>US-10-131-591A-5</th> <th>Sequence 6, Appl1</th>	US-10-131-591A-5	Sequence 6, Appl1
28	217.5	3.8	2659	12 <th>US-10-311-879-28</th> <th>Sequence 28, Appl1</th>	US-10-311-879-28	Sequence 28, Appl1
29	188.5	3.3	2478	9 <th>US-09-815-242-5816</th> <th>Sequence 5816, Ap</th>	US-09-815-242-5816	Sequence 5816, Ap
30	188.5	3.3	2478	9 <th>US-09-815-242-12967</th> <th>Sequence 12967, A</th>	US-09-815-242-12967	Sequence 12967, A
31	187.5	3.3	2368	9 <th>US-09-815-242-5635</th> <th>Sequence 5635, Ap</th>	US-09-815-242-5635	Sequence 5635, Ap
32	187.5	3.3	2368	9 <th>US-09-815-242-12389</th> <th>Sequence 12389, A</th>	US-09-815-242-12389	Sequence 12389, A
33	186.5	3.3	2086	9 <th>US-09-815-242-5639</th> <th>Sequence 5639, Ap</th>	US-09-815-242-5639	Sequence 5639, Ap
34	186.5	3.3	5795	9 <th>US-09-815-242-12610</th> <th>Sequence 12610, A</th>	US-09-815-242-12610	Sequence 12610, A
35	186.5	3.3	6281	9 <th>US-09-815-242-12996</th> <th>Sequence 12996, A</th>	US-09-815-242-12996	Sequence 12996, A
36	185	3.3	1433	10 <th>US-09-801-368-60</th> <th>Sequence 60, Appl1</th>	US-09-801-368-60	Sequence 60, Appl1
37	180	3.2	1029	9 <th>US-09-815-242-5885</th> <th>Sequence 5885, Ap</th>	US-09-815-242-5885	Sequence 5885, Ap
38	180	3.2	1048	9 <th>US-09-815-242-13083</th> <th>Sequence 13083, A</th>	US-09-815-242-13083	Sequence 13083, A
39	177.5	3.1	2434	9 <th>US-09-815-242-5835</th> <th>Sequence 5835, Ap</th>	US-09-815-242-5835	Sequence 5835, Ap
40	176	3.1	3712	14 <th>US-10-108-605-103</th> <th>Sequence 103, App</th>	US-10-108-605-103	Sequence 103, App
41	174	3.1	2437	9 <th>US-09-815-242-5834</th> <th>Sequence 5834, Ap</th>	US-09-815-242-5834	Sequence 5834, Ap
42	167.5	3.0	1069	15 <th>US-10-146-473-77</th> <th>Sequence 77, Appl1</th>	US-10-146-473-77	Sequence 77, Appl1
43	164	2.9	1948	12 <th>US-10-032-585-7611</th> <th>Sequence 7611, Ap</th>	US-10-032-585-7611	Sequence 7611, Ap
44	163.5	2.9	1240	12 <th>US-10-193-764-28</th> <th>Sequence 28, Appl1</th>	US-10-193-764-28	Sequence 28, Appl1
45	163.5	2.9	1226	12 <th>US-10-193-764-26</th> <th>Sequence 26, Appl1</th>	US-10-193-764-26	Sequence 26, Appl1

ALIGNMENTS

RESULT 1

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1

; GENERAL INFORMATION:

; APPLICANT: SAITOH, Shuji

; APPLICANT: TSUZAKI, Yoshinari

; APPLICANT: YANAGIDA, NO. US20010014335A1orU

; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE

; FILE REFERENCE: 981167

; CURRENT APPLICATION NUMBER: US/09/147,052

; CURRENT FILING DATE: 1999-04-05

; PRIOR APPLICATION NUMBER: JP 08-103548

; PRIOR FILING DATE: 1996-03-29

; PRIOR APPLICATION NUMBER: PCT/JP97/01084

; PRIOR FILING DATE: 1997-03-28

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1086

; TYPE: PRT

; ORGANISM: hybrid

US-09-147-052-4

Alignment Scores:

Pred. No.: 0 Length: 1086
 Score: 5619.00 Matches: 1086
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.19% Indels: 0
 DB: 9 Gaps: 0

US-09-147-052-3 (1-3261) x US-09-147-052-4 (1-1086)

Qy	1	ATGCACATATTTAGCGGAATTCATATTTTCCCTATAGTTATTTATGTCACGAAAC	60	301	ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe	320
Db	1	MethIstyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn	20	961	ATGCCCGTGAACCTTCGGCAACGTTTATCATAGTAATACGACTGAGTTTTCACAAATCGC	1020
Qy	61	TCATCTCCGAGTACCCAAATGTGCATCAAGAGAAGTTGTTTCGACGCTCCAGTTGCT	120	321	MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg	340
Db	21	SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer	40	1021	ATCATATAGGACAAATGTTAAACGGCAGGACAGCAACCAATCGGACGACATATTTAGG	1080
Qy	121	GAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACACGTCGGTTCACCGTGATCCGCTA	180	341	IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg	360
Db	41	GluGluGluSerThrPheTyrLeuCysProProProValGlySerThrValIleArgLeu	60	1081	ACAAAATATATGACAGTCACAGTCAAGGTTCAGATGTACAAATATTTCTTCGCTCCGCG	1140
Qy	181	GAACCGCGGGAATTCGCCAACCTAGAAAAGCCACGAGTGGGTGAAGGAATCGCG	240	361	ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly	380
Db	61	GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla	80	1141	GGATTTATGTPAGCATATCAGCCTGTTCTATPCCAAATCCCTGGCTCATATATACCTCAGA	1200
Qy	241	ATATTATTTAAAGAGATATCAGTCCATATATAATTTAAAGTGAGCGCTTTATATAAAT	300	381	GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg	400
Db	81	IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn	100	1201	GAATTGATGAGAGACACAGACCGATGAGATGTCGACCTGGTGAACAAATAAGCATGCA	1260
Qy	301	ATCATTCAGACGACATGCGCGGACGACATATAGACAGATCAGATCATTCGATATACA	360	401	GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla	420
Db	101	IleIleGlnThrThrThrThrThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr	120	1261	ATTTTATAAGAAAAATGCTACCTCATTTGTCAGATTGGCGCCGAGATATTTCGAAATGCACCA	1320
Qy	361	GATAGGACGCGCTTCATTTGAAGAGATCACGGATCTAATTCGACGCAAGAGATGC	420	421	IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro	440
Db	121	AspArgThrProValSerIleGluGluIleThrAspLeuIleAspGlyLysGlyArgCys	140	1321	AATAGAAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTTCATTC	1380
Qy	421	TCATCTAAACAAGATACCTTAGAACAATGTATATGTTGAAGCGTTTGACAGGATGC	480	441	AsnArgLysIleThrLeuAspThrThrAlaIleLysSerThrSerSerValGlnPhe	460
Db	141	SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla	160	1381	GCCATGCTCCAATTTCTTTATGATCATATACAAACCCATATTAATGATGTTTGTAGTAGG	1440
Qy	481	GGAAAAACAAGTACTTCTTAAACCATCAAAATTTCAACGCCCGCAATCTAGGCGATGG	540	461	AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg	480
Db	161	GlyGlyLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp	180	1441	ATTCACACAGCTTGGTGGCAATTCAGAAATAGACAATCTGTTTATGGCAGCAAGGGATA	1500
Qy	541	CACACGACTAATGAGAGCTATACCGTGTGGGATCACCATGGATATATCGAACGGGAAC	600	481	IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle	500
Db	181	HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr	200	1501	AAGATTAACTCTACCGCTACAGCGAGTCGAACATTAGGAAGAGAGTGGTGCAAGATG	1560
Qy	601	TCCGTCAATTTGTATAGTAGGAATAATGGATGCCCGCTCTGTTTCCGTATTCATATTTT	660	501	LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet	520
Db	201	SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe	220	1561	TTGGGGATGTCGCTGCTGATCAGCTGCACCTCATAGATCGGATCGTCACCTTG	1620
Qy	661	GCAATGCCAATGGCGACATCGGAACATATCTCATTTTATGCTCTATCCCGACACAGAG	720	521	LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu	540
Db	221	AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu	240	1621	CAAAATTTCTATGCGAGTTATCACATCCATTAATACATGTTATAGCCGACCATTTGCTCTA	1680
Qy	721	GCTGCCGAGAACCCATGGGATATCCCGAGGATAATTTCAACAACTAGATAGCTATTTT	780	541	GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu	560
Db	241	AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelLysGlnLeuAspSerTyrPhe	260	1681	TTTTTCATATGGAGAAAACCAAGGAACATACAGGAGCAACTCGGTGAAACAACAGAGTTG	1740
Qy	781	TCATGATTTGGACAACCGCTCGAAAGACGCTTCCAGTCAAGCGCTAACTTTCATC	840	561	PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGlnLeuGlyGluAsnAsnGluLeu	580
Db	261	SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle	280	1741	CTTCCAACGCTAGAGGCTGTAGAGCCATGCTCGCTAATCATCATCTAGATATTTCTGTTT	1800
Qy	841	ACATCACATTCACAGTTGGTGGCTGGCTGCCAAACAACTACTCGCTATGTTCAATG	900	581	LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe	600
Db	281	ThrSerHisPheThrValGlyTyrPaspTrpAlaProLysThrThrArgValCysSerMet	300	1801	GGATCCCGTATGCTTTTATTTGAAAACCTATATATTTTGTAAAGATGGTAGAGCTGCCGAT	1860
Qy	901	ACTAAGTGGAAAGAGTGACTGAAATGTTGCGTGCACACAGTTAATGGGAGATACAGATT	960	601	GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp	620
				1861	ATACAGATTCGTACACATTTGTCGAGCTTAATCTAACCCCTGCTAGAGATCGGGAATT	1920
				621	IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle	640
				1921	TTGCTTTATCCGTTTACAAAAAGAGCTGGGTGATGTTGTTGTTATTTGGATTATGCA	1980
				641	LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla	660
				1981	GAAGTACTCGCCCAACTCACTACATGAACCTTAAATTTTATGATACATAAACAAGTAAATA	2040
				661	GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle	680


```
QY 2041 GAAGTGGATCAAAATACGGGGGCTGCAGGAATTCGGCTGTATGTCTATTACTAAAAA 2100
Db 681 GluValAspThrAsnTyrAlaGlyLeuGlnGluPheGlyCysMetSerIleThrLysLys 700
QY 2101 GATCAACAACCAATAATGCGCAACCAATAGCAAGCAGCGGCAATGGAGTTAACAGAT 2160
Db 701 AspAlaAsnProAsnAsnGlyGlnThrGlnLeuGluAlaAlaArgMetGluLeuThrAsp 720
QY 2161 CTAATCAATGCTAAAGCGATGACATATTAGCTTCACATACAAAGACTATGCCAAGATTGAAGCT 2220
Db 721 LeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAla 740
QY 2221 AGTTTATCATCTGCTTATAGTGAAGCTGAACAGCTTACAAATACCTTAATGCAACATTA 2280
Db 741 SerLeuSerSerAlaTyrSerGluAlaGluThrValAsnAsnAsnLeuAlaAsnThrLeu 760
QY 2281 GAACAACCTAAAATGGCTAAACCTAAATTTAGAACTCAGCCATCAACCAAGCTTAATACGAT 2340
Db 761 GluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsnThrAsp 780
QY 2341 AAAACGACTTTGTATATGAACACCCCAATTTAGTTGAAGCATACAAAGCACTAAAAACC 2400
Db 781 LysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeuLysThr 800
QY 2401 ACTTTAGAACACCTGCTACTAACTTGAAGCTTTGTCATCACTCACTGCTTATAATCAAT 2460
Db 801 ThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsnGlnIle 820
QY 2461 CGCAATAATTTAGTGGCTCTATACAAATAAAGCTAGTAGTTTAAATACTAAAAACACTAGAT 2520
Db 821 ArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThrLeuAsp 840
QY 2521 CCACCTAAATGGGGAACCTTTTAGATTCTAATGAGATTACTACAGCTAAATAAGAATATT 2580
Db 841 ProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLysAsnIle 860
QY 2581 AATAATAGCTTATCAACTATTAAATGACAAAAGCTAATGCTGATGCTATCTAATATAGT 2640
Db 861 AsnAsnThrLeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSerAsnSer 880
QY 2641 TTTATTAATAAAGCTGATCAAAATAATCAACAAAGCTTTGTAGGACCTTTTACAACGCT 2700
Db 881 PheIleLysLysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThrAsnAla 900
QY 2701 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGTGCTGTATGATGAACCCGCTCAAT 2760
Db 901 AsnValGlnProSerAsnTyrSerPheValAlaPheSerAlaAspValThrProValAsn 920
QY 2761 TATAAATATGCAAGAGACCGCTTTGGAAATGGTGATGAACCTTCAAGTAGAATTTCTGCA 2820
Db 921 TyrLysTyrAlaArgThrValTrpAsnGlyAspGluProSerSerArgIleLeuAla 940
QY 2821 AACACCAATAGTATCACAGATGTTCTTGGATTTATAGTTAGCTGGGAACAACACGAG 2880
Db 941 AsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsnThrLys 960
QY 2881 TACCAATTTAGTTTGTAGCAACTATGCTCCATCACTGCTTTATTTATTTTCCCTTATAAG 2940
Db 961 TyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 980
QY 2941 TTGGTTAAAGCAGCTGTATGCTTAATAACGTTGGATTGATACAAATACAAATTAATGGAAT 3000
Db 981 LeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGlnTyrLysLeuAsnAsnGlyAsn 1000
QY 3001 GTTCAACAAAGTTGAGTTTGGCACTCAACTAGCTAGTCAAAATATATCTACAGCTTAATCCA 3060
Db 1001 ValGlnGlnValGluPheAlaThrSerThrSerAlaAsnAsnThrThrAlaAsnProThr 1020
QY 3061 CCAGCAGTTGATGAGATTAAGTTGCTAAATTCGTTTTATCAGGTTTAAAGATTTGGCCAA 3120
Db 1021 ProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPheGlyGln 1040
```

```
QY 3121 AACACAATCGAATTAAGTTTCCAAACGGGTGAAGAAATATGATAAAGTTGCCCAATG 3180
Db 1041 AsnThrIleGluLeuSerValProThrGlyAsnMetAsnLysValAlaProMet 1060
QY 3181 ATTGGCAACATTTATCTAGCTCAAAATGAAATATGCTGATAAGATCCCGGGTACCGT 3240
Db 1061 IleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLysIleProGlyTyrArg 1080
QY 3241 CGACCCCGGTACATTTTAA 3258
Db 1081 ArgProGlyThrPheLeu 1086

RESULT 2
US-09-147-052-2
; Sequence 2, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Alignment Scores:
Pred. No.: 2,84e-150 Length: 456
Score: 2015.00 Matches: 396
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.57% Indels: 0
DB: 9 Gaps: 0

US-09-147-052-3 (1-3261) x US-09-147-052-2 (1-456)
QY 2071 GAATTCGGCTGTATGTCTTACTTAAAAAGATGCAAAACCAATAATGGCCAAACCCAA 2130
Db 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
QY 2131 TTAGAAGCAGCGCAATGGAGTTAACAGATCTTAATCAATGCTTAAAGCGATGACATTAGCT 2190
Db 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
QY 2191 TCACCTACAGACTATGCCAAGATTGCAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 2250
Db 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
QY 2251 ACAGTTACAAATAACCTTAAATGCCAACATTAGAACCACTAAAAATGGCTAAAACTAATTTA 2310
Db 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
QY 2311 GAATCAGCATCAACCAAGCTAATACGGATAAAACGACTTTTGTATAATGAACACCAACAA 2370
Db 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAsnGluHisProAsn 160
QY 2371 TTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACCAACGCTGCTACTAACCTGAA 2430
Db 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
QY 2431 GGTTCATCACTCACTGCTTTATTAATCCCAATTAATTTAGTGGATCTATACAAATAA 2490
Db 2491
```

Db 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
QY 2491 GCTAGTACGTTTAACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCT 2550
Db 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
QY 2551 AATGAGATTACTACACTAATAGAAATATTAATATACGTTATCACTTATTAATGAA 2610
Db 221 AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
QY 2611 AAGACTAAATGCTGATGCATTAATAGTCTTATTAATAAAGTGAATCAAAATAATGAA 2670
Db 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGlu 260
QY 2671 CAAAGTTTGTAGGAGCTTTTCAAAACGCTAATGTTTCAACCTTCAAACTACAGTTTGT 2730
Db 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280
QY 2731 GCTTTTAGTGTGATGATACACCCGCTCAATTAATAATATGCAAGAGGACCGTTTGAAT 2790
Db 281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTyrPasn 300
QY 2791 GGTGATGAACCTTCAAGTGAATTCCTGCAACGCTAATGTTTCAACCTTCAAACTACAGTTTGT 2850
Db 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyr 320
QY 2851 ATTATAGTTTACGCTGGACAAACAGCAAGTACCAATTTAGTTTGAACAACTATGCTCCA 2910
Db 321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
QY 2911 TCAACGTTTATTTATTTCCCTTATAAGTTAGTTTAAAGCAGCTGATGCTAATAACGTT 2970
Db 341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
QY 2971 GGATTACAATAAATTAATGGAATGTTTCAACAAGTTGAGTTTCCCACTTCAACT 3030
Db 361 GlyLeuGlyTyrLysLeuAsnAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
QY 3031 AGTGCAAAATTAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3090
Db 381 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys 400
QY 3091 ATCGTTTATCAGTTTAAAGTTTGGCCAAACACAACTCAATTAAGTCTTCCACGGCT 3150
Db 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
QY 3151 GAAGGAAATATGAATAAGTTGCGCAATGATTGGCAACATTTATCTTAGCTCAAAATGAA 3210
Db 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
QY 3211 AATAATGCTGATAAGATCCCGGGTACCGTACCGCGGTACATTTTAA 3258
Db 441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

RESULT 3

US-09-901-572A-3
; Sequence 3, Application US/0901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

Alignment Scores:
Pred. No.: 1,01e-149 Length: 456
Score: 2008.00 Matches: 394
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 35.45% Indels: 0
DB: 12 Gaps: 0
US-09-147-052-3 (1-3261) x US-09-901-572A-3 (1-456)
QY 2071 GAATTCGGCTGTATGCTCTATTACTAAAGATGCAAAACCCAAATTAATGGCCAAACCCAA 2130
Db 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
QY 2131 TTAGAAGCAGCGGAATGGAGTTAACAGATCTAATCAATGCCTAAACGGATGACATAGCT 2190
Db 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
QY 2191 TCACTCAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTCTTATAGTGAAGCTGAA 2250
Db 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
QY 2251 ACAGTTTAACTAACCTTAAATCCACATTAAGACAACTAAATAATGCTAAACTAATTTA 2310
Db 121 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
QY 2311 GAATCAGCCATCAACCAAGCTAATACGGATAAAGACACTTTTGTATTAATGAACCCCAAT 2370
Db 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
QY 2371 TTAGTTGAAGCTATACAAAGCACTAAACACACTTTAGAACACGCTGCTACTAACCTGAA 2430
Db 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
QY 2431 GGTTCGTCATCACTGCTTATTAATCAATTCGCAATTAATTTAGTGGATCTATACATAAA 2490
Db 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200
QY 2491 GCTAGTACGTTTAACTAAACACTAGATCCACTAAATGGGGGAACGCTTTTAGATTCT 2550
Db 201 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 220
QY 2551 AATGAGATTACTACACTAATAGAAATTAATAATACGTTATCACTTATTAATGAA 2610
Db 221 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 240
QY 2611 AAGACTAAATGCTGATGCATTAATAGTCTTATTAATAAAGTGAATCAAAATAATGAA 2670
Db 241 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysValIleGlnAsnAsnGlu 260
QY 2671 CAAAGTTTGTAGGAGCTTTTCAAAACGCTAATGTTTCAACCTTCAAACTACAGTTTGT 2730
Db 261 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 280
QY 2731 GCTTTTAGTGTGATGCTACACCCGCTCAATTAATAATATGCAAGAGGACCGTTTGAAT 2790
Db 281 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTyrPasn 300
QY 2791 GGTGATGAACCTTCAAGTGAATTCCTGCAACGCTAATGTTTCAACCTTCAAACTACAGTTTGT 2850
Db 301 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyr 320
QY 2851 ATTTATAGTTTACGCTGGAAACAAACACGAAAGTACCAATTTAGTTTGAACAACTATGCTCCA 2910
Db 321 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 340
QY 2911 TCAACTGGTTTATTTATTTCCCTTATAAGTTAGTTTAAAGCAGCTGATGCTAATAACGTT 2970
Db 341 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 360
QY 2971 GGATTACAATAAATTAATGGAATGTTTCAACAAGTTGAGTTTCCCACTTCAACT 3030

```
Db 361 GlyLeuGlnTyrLysLeuAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 380
QY 3031 AGTGCATAATACACAGCTAATCAACTCCAGCAGTTGATGAGATTAAAGTCTCTAAA 3090
Db 381 SerAlaAsnThrThrAlaAsnProThrProAlaValAspGluLeuLysValAlaLys 400
QY 3091 ATCGTTTATCAGGTTTAAAGTTGGCCAAACAAATCGAATTAAAGTGTCCACGGGT 3150
Db 401 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 420
QY 3151 GAAGGAATATGAATAAAGTTGGCCCAATGATTTGGCAACATTTATCTTAGTCTCAATGAA 3210
Db 421 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 440
QY 3211 AATAATGCTGATAAGATCCCGGGTACCGTCCAGCCGGTACATTTTA 3258
Db 441 AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456

RESULT 4
US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Alignment Scores:
Pred. No.: 1,48e-143 Length: 384
Score: 1929.50 Matches: 382
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.22% Mismatches: 1
Query Match: 34.06% Indels: 1
DB: 15 Gaps: 1

US-09-147-052-3 (1-3261) x US-10-131-591A-12 (1-384)

QY 2071 GAATTCGGCTGATCTATCTATTACTAAAAAGATGCAACCCCAATAAATGGCCAAACCCAA 2130
Db 1 GluPhe---CysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 19
QY 2131 TTAGAAGCAGCGCAATGGAGTTAAACAGATCTAACTCAATGCTAAAGCGATGACATTAGCT 2190
Db 20 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 39
QY 2191 TCATCACAAGACTATGCCAAGATGCAAGCTAGTTATCATCTCTATGCTATAGTGAAGCTGAA 2250
Db 40 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 59
QY 2251 ACAGTTAACAATACTTAATGCCACATAGACAACTAAATAATGGTAAACTAATTTA 2310
Db 60 ThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 79
QY 2311 GAATCAGGCATCAACCAAGCTAATACGGATAAAGACGCTTTTGTAAATGAACACCAAAAT 2370
Db 80 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 99
QY 2371 TTAGTTGAAGCATCAACAGCACTAAAAACCACTTTAGAACACAGCTGCTACTAACCTTGA 2430
Db 100 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 119
QY 2431 GGTTCATCATCACTGCTTTATTAATCAAAATTCGAATAAATTTAGTGGATCTATACATAAA 2490
```

```
Db 120 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 139
QY 2491 GCTAGTACTTTAATAACTAAACACTAGATCCACTAAATGGGGAACGGCTTTTAGATTCT 2550
Db 140 AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer 159
QY 2551 AATGAGATTACACAGCTAATAAGAAATATTAATAATACGTTATCAACTATTATTAACAA 2610
Db 160 AsnGluIleThrThrValAsnArgAsnIleAsnAsnThrLeuSerThrIleAsnGluGln 179
QY 2611 RAGACTAATGCTGATGCATTATCTAATAGTTTATTAATAAAGTATTCACAAATATGAA 2670
Db 180 LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu 199
QY 2671 CAAAGTCTTTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAACCTACACAGTTTGT 2730
Db 200 GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal 219
QY 2731 GCTTTTAGTGTGATGTAAACCCGTCATTAATAATATGAAGAAGGACCGTTTGGAAAT 2790
Db 220 AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaArgThrValTyrAsn 239
QY 2791 GGTGATGACCTTCAAGTAGAATTTCTGCAACACAGCAATAGTATCACAGATGTTCTGG 2850
Db 240 GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTyr 259
QY 2851 ATTTATAGTTTACCTGGCAACAAACACAGAGTAGTACCATTGTTTGGTAAAGCAGCT 2910
Db 260 IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro 279
QY 2911 TCACTGGTTATTTATATTTCCCTTATAGTTGGTAAAGCAGCTGATCTAATACGTT 2970
Db 280 SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnVal 299
QY 2971 CGATTACATCAAAATTAATAATGGAATGTTCAACAGTTGAGTTTCCCACTTCAACT 3030
Db 300 GlyLeuGlnTyrLysLeuAsnGlyAsnValGlnGlnValGluPheAlaThrSerThr 319
QY 3031 AGTCAAAATAATACATACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAA 3090
Db 320 SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluLeuLysValAlaLys 339
QY 3091 ATCGTTTATCAGGTTTAAAGTTGGCCAAACAAACAAATCGAATTAAGTGTTCACACGGT 3150
Db 340 IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly 359
QY 3151 GAAGCAATATCAATAAAGTTGCGCAATGATGCGCAATGATGCGCAATTTATCTTACCAATGAA 3210
Db 360 GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu 379
QY 3211 AATAATGCTGATAAG 3225
Db 380 AsnAsnAlaAspLys 384

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of
```

OTHER INFORMATION: pn240K-S
US-10-131-591A-23

Alignment Scores:

Pred. No.: 1,51e-131 Length: 357
Score: 1777.00 Matches: 352
Percent Similarity: 98.60% Conservative: 0
Best Local Similarity: 98.60% Mismatches: 5
Query Match: 31.37% Indels: 0
DB: 15 Gaps: 0

US-09-147-052-3 (1-3261) x US-10-131-591A-23 (1-357)

```
QY 2155 ACAGATCTAATCAATGCTAAAGCGATGACATTAAGCTTCACATCAAGACTATGCCAAGATT 2214
Db 1 ThrAspLeuLeuAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIle 20
QY 2215 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGCTTAACATAACCTTAAGTCA 2274
Db 21 GluAlaSerLeuSerAlaTyrSerGluAlaGluThrValAsnAsnLeuGlnAla 40
QY 2275 ACATTAGAACAACTAAATGGCTAAACTAAATTTAGAAATCAGCCATCAACCAAGCTAAT 2334
Db 41 ThrLeuGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsn 60
QY 2335 ACGATAAAACGACTTTTGATATGAACACCCCAATTTAGTTGAAGCATACAAGCACTA 2394
Db 61 ThrAspLysThrThrPheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeu 80
QY 2395 AAAACCACTTTAGAACACGCTGCTACTTAACCTTTGAAGGTTTGTCATCAACTGCTTATAAT 2454
Db 81 LysThrThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerThrAlaTyrAsn 100
QY 2455 CAATTCGCGCAATTAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAATACTAAACA 2514
Db 101 GlnIleArgAsnAsnLeuValAspLeuLysThrAsnLysAlaSerLeuIleThrLysThr 120
QY 2515 CTAGATCCACTAAATGGGGAACGCTTTTAGATTCATTAATGAGATTACTACAGCTAATAAG 2574
Db 121 LeuAspProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLys 140
QY 2575 AATATTAAATACGTTATCAACTATTAAATGAACAAAGACTAATGCTGATGCTATATCT 2634
Db 141 AsnIleGlnAsnThrLeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSer 160
QY 2635 AATAGTTTTATTAATAAAGTATCAATAATAATGACAAAGTTTGTAGGACTTTTACA 2694
Db 161 AsnSerPheIleLysLysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThr 180
QY 2695 AAGCGTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGTGCTGATGTAACACCC 2754
Db 181 AsnAlaAsnValGlnProSerGlnTyrSerPheValAlaPheSerAlaAspValThrPro 200
QY 2755 GTCAATTATATATGCAAGAAGACCGTTTGAATGGTGATCAACCTTCAAGTAGAATT 2814
Db 201 ValAsnTyrLysTyrAlaArgThrValThrAsnGlyAspGluProSerSerArgIle 220
QY 2815 CTTGCAAAACAGAAATAGTATACAGATGTTTCTTGGATTTATAGTTAGCTGGAACAAAC 2874
Db 221 LeuAlaAsnThrAsnSerIleThrAspValSerTyrIleTyrSerLeuAlaGlyThrAsn 240
QY 2875 ACGAAGTACCAATTTAGTTTACCACTATGGTCCCACTGTTTATTTATTTTCCCT 2934
Db 241 ThrLysTyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPhePro 260
QY 2935 TATAAGTTGGTTAAAGCAGCTGATGCTAATAAGCTTGGATTACAATAACAATAATAAT 2994
Db 261 TyrLysLeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGlnTyrLysLeuAsnAsn 280
QY 2995 GGAATGTTCAACAGTTAGTTGGCCACTTCAACTAGTGTCAAAATAATACTACAGCTAAT 3054
Db 281 GlyAsnValGlnGlnValGluPheAlaThrSerThrSerAlaGlnGlnThrThrAlaAsn 300
```

```
QY 3055 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTTATCAGGTTTAAGATT 3114
Db 301 ProThrProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPhe 320
QY 3115 GGCACAAACACAATCGAATTAAGTTTCCAAACGGGTCAAGGAAATATGATAAAGTTGCG 3174
Db 321 GlyGlnAsnThrIleGluLeuSerValProThrGlyGluGlyAsnMetAsnLysValAla 340
QY 3175 CCAATGATTGGCAACATTTATCTTTAGCTCAAAATGAAAATATGCTGATAAG 3225
Db 341 ProMetIleGlyAsnIleTyrLeuSerSerAsnGluAsnAlaAspLys 357

RESULT 6
US-09-894-998-18
; Sequence 18, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITL OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-09-894-998-18

Alignment Scores:
Pred. No.: 3,41e-128 Length: 904
Score: 1736.00 Matches: 363
Percent Similarity: 59.83% Conservative: 142
Best Local Similarity: 43.01% Mismatches: 269
Query Match: 30.64% Indels: 70
DB: 9 Gaps: 16

US-09-147-052-3 (1-3261) x US-09-894-998-18 (1-904)
QY 58 AACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTG 117
Db 82 AsnAlaThrValAlaAlaGlyHisAlaThrLeuArgAlaHisLeuArgGluIleLysVal 101
QY 118 TCTGAGGAAGAGTCTAGCTTTTATCTTTGTCCTCCACCCAGTGGTTCAACCGTATCCGT 177
Db 102 GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValValGln 121
QY 178 CTAGACCGCGCGGAAATGTCCTCGAACCTAGAAAGCCAGCCGAGTGGGTGAAGGAATC 237
Db 122 PheGluGlnProArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyIle 141
QY 238 GCGATATTATTAAAGAGAATATCAGTCCATATAAATTTAAAGTGACGCTTTATTATATAA 297
Db 142 AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 161
QY 298 AATATCATTCAGACGACATCGACGGGACGACATATAGACAGATCACTAATCATATAT 357
Db 162 AspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 181
QY 358 ACAGATAGGACGCCGTTTCCATTTGAGATCATCAGGATCATTCGACGCAAGGAAGA 417
Db 182 GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnThrLysGlyVal 201
QY 418 TGTCTCATCTAAACAGATACCTTAGAAACAATGTATATCTTGAAGCGTTTTCACAGGAT 477
Db 202 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgAsp 221
```

478	QC	CGCGGAGAAAAACAGATCTTCTTAAACACCATCAAAATTCACACGCGCCGAATCTAGGCCA	537
222	DB	AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArgGly	241
538	QC	TGCGACAGCACTAATGAGACGTATACCGTGTGGGGATCACCATGGATATATCATCCGGA	597
242	DB	TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly	261
598	QC	ACCTCCGCTCAATTTGTATAGTAGAGAAATGCATCCCGCTCTGTTTCCGTTATCATAT	657
262	DB	ThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspGlu	281
658	QC	TTTGCAATGCCAATGGCGACATCGCGAACATATCTCCATTTATGGTCTATCCCAACCA	717
282	DB	PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg--Glu	300
718	QC	GAGGCTCCGCCAAGACCCATGGGATATCCCAAGATAATTTCAACCACTAGATAGCTAT	777
301	DB	GlySerHisThrGluHisThrSerTyrAlaAlaAspArgPheLysGlnValAspGlyPhe	320
778	QC	TTTTCATGAGTGGCAAGCGCTCGAAAAAGCGCTTCACGTCAAGCGCTTAATCTTCTC	837
321	DB	TyrAlaArgAspLeuThrThrLysAlaArgAlaThrSerProThrThrArgAsnLeuLeu	340
838	QC	ATCACATCACATTCACAGTTGGGTGGGACTGGGCTCCCAAAACTACTCTGTATGTCTCA	897
341	DB	ThrThrProLysPheThrValAlaTirPaspTirPValProLysArgProAlaValCysThr	360
898	QC	ATCAGTAAGTGGAAAGAGGTGACTGTAATCTTCGGTGCACACAGTTAATGGGAGATACA	957
361	DB	MetThrLysTrpGlnGluValAspGluMetLeuArgAlaGluTyrGlyGlySerPheArg	380
958	QC	TTTATGGCCGCTGAATCTTCGCAAGCTTTATCAGTAATACGACTAGTTTGATCCAAAT	1017
381	DB	PheSerSerAspAlaIleSerThrThrPheThrThrAsnLeuThrGluTyrSerLeuSer	400
1018	QC	CGCATCATATTAGCAATGATTTAAACGGCGAGCAGACAGCAATCGAGCAGATATTT	1077
401	DB	ArgValAspLeuGlyAspCysIleGlyArgAspAlaArgGluAlaIleAspArgMetPhe	420
1078	QC	AGGACAAAATATATGACAGTCAAGTCAAGTTCAGATGATGACATATTTCTTTGGCTCTC	1137
421	DB	AlaArgLysTyrAsnAlaThrHisIleLysValGlyGlnProGlnTyrTyrLeuAlaThr	440
1138	QC	GGGGATTTATTTAGCATCATCGCTGTCTTATCCAAATCCCTGGCTCATATGTACTCT	1197
441	DB	GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnThrLeuAlaGluLeuTyrVal	460
1198	QC	AGAAATTGATGAGACACACAGACCGATGAGATGCTCGACCTGTTAAACAATAAGCAT	1257
461	DB	ArgGluTyrMetArgGluGln-----AspArgLys--	470
1258	QC	GCAATTTAAGAAAAATGCTACTCTACTTCAGGATTCGGCGGAGATATCGAAATGCA	1317
471	DB	-----ProArgAsnAlaThrPro-----AlaProLeuArgGluAla	482
1318	QC	CCAAATAGAAAAATAACATTTAGACGACACACACAGCTATTAAATCGACATCGTGTCAA	1377
483	DB	ProSerAlaAsnAlaSerValGlu-----ArgIleLysThrThrSerSerIleGlu	499
1378	QC	TTCCCATGCTCCAATTTCTTTATGATCATATACAACCCATATTAATGATATGTTTAGT	1437
500	DB	PheAlaArgLeuGlnPheThrTyrAsnHisIleGlnArgHisValAsnAspMetLeuGly	519
1438	QC	AGAAATGCCACACGTTGGTCGAATTCGCAAAATAGAGAAGCTGTTTATGGCAGGAGGG	1497
520	DB	ArgIleAlaValAlaTrpCysGluLeuGlnAsnHisGluLeuThrLeuTrpAsnGluAla	539
1498	QC	ATAAGATTAATCCTAGCGCTACAGCGAGTGCACATATAGGAAGGACAGTGGCTGCAAG	1557
540	DB	ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgValSerAlaArg	559
1558	QC	ATGTTGGGGATGTCGCTGCTGTATCGAGCTGCATGCTATAGATCGGAATCCGTCACCT	1617

Db	560	MetLeuGlyAspValMetAlaValSerThrCysValProValAlaLapProAspAsnValIle	579
QY	1618	TTGCAGAAATTCATGCGCAGTTATACATACATACATGTTATATACCCACCATTCGGTT	1677
Db	580	ValGlnAsnSerMetArgValSerSerArgProGlyThrCysTyrSerArgProLeuVal	599
QY	1678	CTATTTTCATATGAGAGAAACCAAGGAAAC---ATACAGGAGCAACTCGGTGAAACAAC	1734
Db	600	SerPheArgTyr---GluAspGlnGlyProLeuIleGluGlyGlnLeuGlyGluAsnAsn	618
QY	1735	GAGTTGCTTCCAGCGCTAGAGCGTGTAGAGCGATGCTCGGCTAATCATCTAGATATTTT	1794
Db	619	GluLeuArgLeuThrArgAspAlaLeuGluProCysThrValGlyHisArgArgTyrPhe	638
QY	1795	CTGTTTGTGATCCGGTTATGCTTTATTTGAAACATATAATTTGCTTAAAGATGGTAGACGCT	1854
Db	639	IlePheGlyGlyTyrValTyrPheGluGluTyrAlaTyrSerHisGlnLeuSerArg	658
QY	1855	GCCGATATACAGATTGCTAGCACATTTGTCGAGCTTAATCATCCCTGCTAGAGATCGG	1914
Db	659	AlaAspValThrThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAspHis	678
QY	1915	GAATTTTGCTTTATCCGTTTACACAAAGACAGATGGGTGATGTTGGTGTATTCGAT	1974
Db	679	GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp	698
QY	1975	TATGCAGAGTAGTCGCGCAATCAACTACATCACTTAAATTTTATGACATAAACAAA	2034
Db	699	TyrThrGluValGlnArgArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr	718
QY	2035	GTAATAGAGTGGATACAAAT-----TACGCGGGGTGCAGGAATTC-----	2076
Db	719	ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPhePheGluGly	738
QY	2077	-----GGCTGTATGCTTATTACTTAAAAAAGATGCA	2106
Db	739	MetGlyAspLeuGlyArgAlaValGlyLysValValMetGlyValValGlyGlyValVal	758
QY	2107	AACCCAATAATAGCCAAACCAATTAGAAGCAGCGCGAATGGAGTTAAACAGATCTAATC	2166
Db	759	SerAlaValSerGlyValSerSerPhe-----MetSerAsnProPhe	772
QY	2167	AATGCTAAAGCGATGACATTAGCTTCTACTACAAGACTATGCCAGATTGAAGCTGTTA	2226
Db	773	GlyAlaLeuAlaValGlyLeuLeuValLeu-----AlaGlyLeu	785
QY	2227	TCATCTGCTTATAGTAGAGCTGAACAGTT-----AACATAACCTTTAATGCA	2274
Db	786	ValAlaAlaPhePheAlaPheArgTyrValLeuGlnLeuGlnArgAsnProMetLysAla	805
QY	2275	-----ACATTAGAACAACTAAATAATGGCTAAACATAATTTAGAATCAGCCATC	2322
Db	806	LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyGlyValGlyGlyGlu	825
QY	2323	AACCAAGCTAATACGGATAAACACGCTTTTGTATTAATGAACACCCAAATTTAGTTGAACGA	2382
Db	826	GlyGluGluGlyAlaGluGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla	843
QY	2383	-----TACAAAGCACTAAACACCTTTAGAA-----CAACGTGCTACT	2421
Db	844	ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla	863
QY	2422	AACCTTGAAGGTTTGTTCATCACTGCTTATAATCAAAATTCGCAATAATTTAGTGCATCTA	2481
Db	864	ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys	883
QY	2482	TACAATAAGCT	2493
Db	884	ArgAsnLysAla	887

Db 639 IlePheGlyGlyTyrValTyrPheGluGluTyrAlaTyrSerHisGlnLeuSerArg 658
QY 1855 GCCGATATACAGATTGCTAGACATTTGTCAGCTTAATCTAACCCCTGCTAGAAAGTCGG 1914
Db 659 AlaAspValThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAspHis 678
QY 1915 GAAATTTGCGCTTATCCGTTTACACAAAGAGATTCCCGATGTTGGTGTATGGAT 1974
Db 679 GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp 698
QY 1975 TATGCAAGTAGTACGCGCCGAATCAACTACATCAACTTAAATTTTATGACATAACAAA 2034
Db 699 TyrThrGluValGlnArgArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr 718
QY 2035 GTAATGAGAGTGGATACAAAT-----TACGCGGGCTCGCAAGATTC----- 2076
Db 719 ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPheGluGly 738
QY 2077 -----GGCTGTATGCTCTATTACTAAAAAAGATGCA 2106
Db 739 MetGlyAspLeuGlyArgAlaValGlyLysValValMetGlyValValGlyValVal 758
QY 2107 AACCCAAATAATGGCCAAACCAATTAGAAGCAGCGGAATGGAGTTACAGATCTAATC 2166
Db 759 SerAlaValSerGlyValSerPhe-----MetSerAsnProPhe 772
QY 2167 AATGCTAAGCGATGACATAGCTTCACTACAGACTATGCCAAGATTGAAGCTAGTTTA 2226
Db 773 GlyAlaLeuAlaValGlyLeuLeuValLeu-----AlaGlyLeu 785
QY 2227 TCATCTGCTTATAGTGAAGCTGAACAGTT-----AACAAATACCTTAATGCA 2274
Db 786 ValAlaAlaPhePheAlaPheArgTyrValLeuGlnLeuGlnArgAsnProMetLysAla 805
QY 2275 -----ACATTAGAACAACTAAATAATGGCTAAACTAAATTTAGAACTCAGCCATC 2322
Db 806 LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyValGlyGlyGlu 825
QY 2323 AACCAAGCTAATACCGATAAAGCACTTTGTATATGATGAACACCCAAATTTAGTGAAGCA 2382
Db 826 GlyGluGluGlyAlaGluGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
QY 2383 -----TACAAACGACTAAACCACTTTAGAA-----CAACGTGCTACT 2421
Db 844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
QY 2422 AACCTTGAAGTTTGTCACTCAACTGCTTATAATCAAAATTCGCAATAATTTAGTGGATCTA 2481
Db 864 ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys 883
QY 2482 TACATAAAGCT 2493
Db 884 ArgAsnLysAla 887

RESULT 8

US-10-237-551-18

; Sequence 18, Application US/10237551

; Publication No. US20030165820A1

; GENERAL INFORMATION:

; APPLICANT: Day, Craig H.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Parsons, Joseph M.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION

; FILE REFERENCE: 210121.538C3

; CURRENT APPLICATION NUMBER: US/10/237,551

; CURRENT FILING DATE: 2002-09-06

; NUMBER OF SEQ ID NOS: 254

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 18

; LENGTH: 904

; TYPE: PRT

; ORGANISM: Herpes simplex virus

US-10-237-551-18

Alignment Scores:

Pred. No.: 3,41e-128 Length: 904
Score: 1736.00 Matches: 363
Percent Similarity: 59.83% Conservative: 142
Best Local Similarity: 43.01% Mismatches: 269
Query Match: 30.64% Indels: 70
DB: 12 Gaps: 16

US-09-147-052-3 (1-3261) x US-10-237-551-18 (1-904)

QY 58 AACTCATCTCCAGAGTACCCAAAATGTGCATCAAGAGAAGTTGTTTCGAGCGTCCAGTTG 117
Db 82 AsnAlaThrValAlaAlaGlyHisAlaThrLeuArgAlaHisLeuArgGluIleLysVal 101
QY 118 TCTGAGGAAGAGTCTACGTTTATCTTCTGTCCTCCACAGTGGGTTCACACCGTATCCGT 177
Db 102 GluAsnAlaAspAlaGlnPheTyrValCysProProThrGlyAlaThrValValGln 121
QY 178 CTAGAACCCGCGCAAAATGTCCGAACTAGAAAAGCCAGAGTGGGTGGAAGGAATC 237
Db 122 PheGluGlnProArgArgCysProThrArgProGluGlyGlnAsnTyrThrGluGlyIle 141
QY 238 GCGATATTATTAAAGAGAAATATCAGTCCATATAAATTTAAAGTACGCTTATTATATAA 297
Db 142 AlaValValPheLysGluAsnIleAlaProTyrLysPheLysAlaThrMetTyrTyrLys 161
QY 298 AATATCATTACAGACGACATGAGCGGGGAGCAGACATATAGACATCACTAATCGATAT 357
Db 162 AspValThrValSerGlnValTyrPheGlyHisArgTyrSerGlnPheMetGlyIlePhe 181
QY 358 ACAGATAGACGCCCGCTTTCCATTGAAGAGATCAGCATCTAATCAGCGCAAGGAAGA 417
Db 182 GluAspArgAlaProValProPheGluGluValIleAspLysIleAsnThrLysGlyVal 201
QY 418 TCGTCTCATCTAAAGCAGATACCTTACAAACAATGTATATGTGAACCGTTTGACAGGAT 477
Db 202 CysArgSerThrAlaLysTyrValArgAsnAsnMetGluThrThrAlaPheHisArgAsp 221
QY 478 GCGGGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCTAGGGCA 537
Db 222 AspHisGluThrAspMetGluLeuLysProAlaLysValAlaThrArgThrSerArgGly 241
QY 538 TGGCACACGACTAATGAGACGTATACCGTGTGGGGATCACCATGGATATATCGAACGGGA 597
Db 242 TrpHisThrThrAspLeuLysTyrAsnProSerArgValGluAlaPheHisArgTyrGly 261
QY 598 ACCTCGCTCAATGTATAGTAGAGGAATGGATGCCCGCTCTGCTGTTCCGTATTCATAT 657
Db 262 ThrThrValAsnCysIleValGluGluValAspAlaArgSerValTyrProTyrAspGlu 281
QY 658 TTTGCAATGGCCAATGGCAGCATCGGCAACATATCTCCATTTTATGGTCTATCCCCACCA 717
Db 282 PheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg---Glu 300
QY 718 GAGGCTGCCGACAGAACCCATGGGATATCCCCAGGATAATTTCAACAACACTAGATAGTAT 777
Db 301 GlySerHisThrGluHisThrSerTyrAlaAlaAspArgPheLysGlnValAspGlyPhe 320
QY 778 TTTTCAATGGATTGGACAGCGCTCGAAAAGCAAGCGTTCACGCTCAAGCGTAACCTTCTC 837
Db 321 TyrAlaArgAspLeuThrThrLysAlaArgAlaThrSerProThrThrArgAsnLeuLeu 340
QY 838 ATCATATCACACTTTCACAGTGGGTGGGCTGGGCTCCAAAACACTACTCGTGTATCTTCA 897
Db 341 ThrThrProLysPheThrValAlaThrValPaspThrPvalProLysArgProAlaValCysThr 360
QY 898 ATGACTAAGTGGAAAGAGGTGACTGAAATGTTGCGTGCACACAGTAAATGAGGATACAGA 957
Db 361 MetThrLysTrpGlnGluValAspGluMetLeuArgAlaGluTyrGlySerPheArg 380
QY 958 TTTATGGCGCGTGAACCTTTTCGGCAACGTTTTATCAGTAATACGACTGAGTTTGATCCAAAT 1017


```

Db 381 PheSerAspAlaIleSerThrThrPheThrAsnLeuThrGluTyrSerLeuSer 400
QY 1018 CGCATATATAGACATATATTAACGCGAGCAGACGAGCAATCGAGAGATTTT 1077
Db 401 ArgValAspLeuGlyAspCysIleGlyArgAspAlaArgGluAlaIleAspArgMetPhe 420
QY 1078 AGGCAAAATATATGACATCAGCTCAGTCAAGTGGACATGATACAATTTCTTGGCTCTC 1137
Db 421 AlaArgLysTyrAsnAlaThrHisIleLysValGlyGlnProGlnTyrTyrLeuAlaThr 440
QY 1138 GGGGGATTTATGAGCATATCAGCCCTGTTCTATCAAAATCCCTGGCTATATGTACTC 1197
Db 441 GlyGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnThrLeuAlaGluLeuTyrVal 460
QY 1198 AGAAATGATGAGACACACAGCAGCAGATGCTCGACCTGGTGAACATAATGACAT 1257
Db 461 ArgGluTyrMetArgGluGln-----AspArgLys--- 470
QY 1258 GCATTTATAGAAAATGCTACTCTATGTCAGGATGCGGGCAGATATTCGAATGCA 1317
Db 471 -----ProArgAsnAlaThrPro-----AlaProLeuArgGluAla 482
QY 1318 CCAATAGAAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTGTCAA 1377
Db 483 ProSerAlaAsnAlaSerValGlu-----ArgIleLysThrThrSerIleGlu 499
QY 1378 TTCGCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATTAATGATATGTTAGT 1437
Db 500 PheAlaArgLeuGlnPheThrTyrAsnHisIleGlnArgHisValAsnAspMetLeuGly 519
QY 1438 AGGATGCCACAGCTTGGTGGAAATGCCAAGATAGAACTGTTTATGGCAGGAGGG 1497
Db 520 ArgIleAlaValAlaTyrCysGluLeuGlnAsnHisGluLeuThrLeuTyrAsnGluAla 539
QY 1498 ATAAGATTAATCTACGCTACAGCAGTGCACATTTAGAGAGAGTGGCTGCAAG 1557
Db 540 ArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgValSerAlaArg 559
QY 1558 ATGTTGGGGATGTCGCTGCTGATCGAGTGCCTACTATAGATCGGAAATCGCTCACT 1617
Db 560 MetLeuGlyAspValMetAlaValSerThrCysValProValAlaProAspAsnValIle 579
QY 1618 TTGCATAATCTATGCGGATATCACATCCACTAATACATGTTATACCGCAGCATGCTT 1677
Db 580 ValGlnAsnSerMetArgValSerSerArgProGlyThrCysTyrSerArgProLeuVal 599
QY 1678 CTATTTTCATATGAGAAAACCAAGGAAAC---ATACAGGACACAACTCGGTGAAACAAAC 1734
Db 600 SerPheArgTyr---GluAspGlnGlyProLeuIleGluGlnLeuGlyGluAsnAsn 618
QY 1735 GAGTTGCTTCAACGCTAGAGCTGTAGAGCCATGCTCGGCTAATCATCCTAGATATTTT 1794
Db 619 GluLeuArgLeuThrArgAspAlaLeuGluProCysThrValGlyHisArgTyrPhe 638
QY 1795 CTGTTTGGATCGGTTATGCTTTATTTGAAACTATATTTTGTAGATGGTAGAGCT 1854
Db 639 IlePheGlyGlyIleValTyrValTyrPheGluTyrAlaTyrSerHisGlnLeuSerArg 658
QY 1855 GCCGATATACAGATGCTAGCACATTTGTCGAGCTTAATCTAACCCCTAGAGATCGG 1914
Db 659 AlaAspValThrThrValSerThrPheIleAspLeuAsnIleThrMetLeuGluAsnHis 678
QY 1915 GAAATTTTCCGTTTATCCGTTTACAAAGAGTGGCTGATGTTGCTGTTATGATGAT 1974
Db 679 GluPheValProLeuGluValTyrThrArgHisGluIleLysAspSerGlyLeuLeuAsp 698
QY 1975 TATGACAGATGAGTCCCGGCAATCACTACATCACTAAATTTTATGACATAACAAA 2034
Db 699 TyrThrGluValGlnArgArgAsnGlnLeuHisAspLeuArgPheAlaAspIleAspThr 718
QY 2035 GTAATAGAGTGGATACAAAT-----TACCGGGGCTGCGAGGAATTC----- 2076

```

```

Db 719 ValIleArgAlaAspAlaAsnAlaAlaMetPheAlaGlyLeuCysAlaPhePheGluGly 738
QY 2077 -----GGCTGATGCTCTATTACTATAAAAAAGATGCA 2106
Db 739 MetGlyAspLeuGlyArgAlaValGlyLysValValMetGlyValValGlyGlyValVal 758
QY 2107 AACCCAAATTAATGGCCAAACCCCAATTAGAAGCAGCGGAATGAGTAAACAGATCTAATC 2166
Db 759 SerAlaValSerGlyValSerSerPhe-----MetSerAsnProPhe 772
QY 2167 AATGCTAAAGCGATGACATAGCTTCACTACAAGACTATGCCAAGATTGAAGTAGTTTA 2226
Db 773 GlyAlaLeuAlaValGlyLeuLeuValLeu-----AlaGlyLeu 795
QY 2227 TCATCTGCTTATAGTGAAGCTGAACAGTT-----AACATAACCTTAATGCA 2274
Db 786 ValAlaAlaPheAlaPheArgTyrValValGlnLeuGlnArgAsnProMetLysAla 805
QY 2275 -----ACATTAGAACAACTAAATAATGGCTAAATACTAATTTAGTAATCAGCATC 2322
Db 806 LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyGlyValGlyGlyGlu 825
QY 2323 AACCAAGCTAATACGGATAAAAGCACTTTTGATAATGAACACCCAAATTTTGTGAAGCA 2382
Db 826 GlyGluGluGlyAlaGluGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
QY 2383 -----TACAAAGCACTAAACCACTTTAGAA-----CAACGTGCTACT 2421
Db 844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
QY 2422 AACCTTGAAGTTGTCATCACTGCTTATTAATCAATTCGCAATATTTAGTGGATCTA 2481
Db 864 ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys 883
QY 2482 TACAATAAAGCT 2493
Db 884 ArgAsnLysAla 887

```

RESULT 9

```

US-10-121-988-18
; Sequence 18, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McCowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121.988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

```

```

Alignment Scores:
Pred. No.: 3,41e-128 Length: 904
Score: 1736.00 Matches: 363
Percent Similarity: 59.83% Conservative: 142
Best Local Similarity: 43.01% Mismatches: 269
Query Match: 30.64% Indels: 70
DB: 16
US-09-147-052-3 (1-3261) x US-10-121-988-18 (1-904)

```



```

Db      773 GlyAlaLeuAlaValGlyLeuValLeu-----AlaGlyLeu 785
QY      2227 TCATCTGCTTATAGTGAACCTGAACAGTT-----AACATAACCTTATGCA 2274
Db      786 ValAlaAlaPhePheAlaPheArgTyrValLeuGlnLeuGlnArgAsnPrometLysAla 805
QY      2275 -----ACATTAGAACAACTAAATAATGGCTAAATACTAATTAGAAATCAGCCATC 2322
Db      806 LeuTyrProLeuThrThrLysGluLeuLysThrSerAspProGlyGlyValGlyGlu 825
QY      2323 AACCAAGCTAATACGAGTAAACAGCTTTTGTATATGAATGAACACCCAAATTTAGTTGAAGCA 2382
Db      826 GlyGluGluGlyAlaGluGlyGlyGlyPheAsp-----GluAlaLysLeuAlaGluAla 843
QY      2383 -----TACAAAGCACTAAATAACACCTTTAGAA-----CAACGTGCTACT 2421
Db      844 ArgGluMetIleArgTyrMetAlaLeuValSerAlaMetGluArgThrGluHisLysAla 863
QY      2422 AACCTTGAAGTTGTTCATCACTGCTTATATCAATCAATTCGCAATAATTAGTGGATCTA 2481
Db      864 ArgLysLysGlyThrSerAlaLeuLeuSerSerLysValThrAsnMetValLeuArgLys 883
QY      2482 TACAATAAAGCT 2493
Db      884 ArgAsnLysAla 887

RESULT 10
US-10-223-538-2
; Sequence 2, Application US/10223538
; Publication No. US20030120060A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Karl, Csaba
; TITLE OF INVENTION: No. US20030120060A1el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/223,538
; FILING DATE: 19-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

```

US-10-223-538-2
Alignment Scores:
Pred. No.: 1,01e-52 Length: 907
Score: 778.50 Matches: 255
Percent Similarity: 43.68% Conservative: 153
Best Local Similarity: 27.30% Mismatches: 353
Query Match: 13.74% Indels: 173
DB: 15 Gaps: 29

US-09-147-052-3 (1-3261) x US-10-223-538-2 (1-907)
QY      22 TGCATATTTTCTTATAGTATATCTATATGGTACGAACCTCATCTCGAGTACC----- 75
Db      11 CysValAsnLeuCysIleValCysLeuGlyAlaAlaValSerSerSerThrArgGly 30
QY      75 -----
Db      31 ThrSerAlaThrHisSerHisHisSerSerHisThrThrSerAlaAlaHisSerArgSer 50
QY      76 -----CAAAATGTGACATCAAGAGAGTTGTTTCGAGC----- 108
Db      51 GlySerValSerGlnArgValThrSerSerGlnThrValSerHisGlyValAsnGluThr 70
QY      109 -----GTCAGTTGCTGAGGAGAGTCTACG 135
Db      71 IleTyrAsnThrThrLeuLysTyrGlyAspValValGlyValAsnThrThrLysTyrPro 90
QY      136 TTTTATCTTTGTCCTCCACAGCAGTTCACCGTTCACCGTTCGCTAGAACCGCGCGAATA 195
Db      91 TyrArgValCysSerMetAlaGlnGlyThrAspLeuIleArgPheGluArgAsnIleVal 110
QY      196 TGTCGCGAACCTAGAAAAGCCACCGAG---TGGGGTGAAGGAATCGCGATATTATTAAA 252
Db      111 CysThrSerMetLysProIleAsnGluAspLeuAspGluGlyIleMetValValTyrLys 130
QY      253 GAGAATATACAGTCCATATAATTAAGTACGCTTTATATATAATAATATCATT----- 306
Db      131 ArgAsnIleValAlaHisThrPheLysValArgValTyrGlnLysValLeuThrPheArg 150
QY      307 CAGACGACGACATGCGGCGGACGACATATAGACAGATCACTACTCATCATACATACAGATAGG 366
Db      151 ArgSerTyrAlaTyrIleHisThrThrTyrLeuLeuGlySerAsn-----ThrGluTyr 168
QY      367 ACCCGCGTTTCCATTCGAAGAGATCAGGATCTAATCGACGGCAAGAGATGC---TCA 423
Db      169 ValAlaProProMetTyrGluIleHisHis---IleAsnSerHisSerGlnCysTyrSer 187
QY      424 TCTAAAGCAAGATACCTTAGAAAACATGTATATGTTGACGCGTTTGACAGGGATCGGGA 483
Db      188 SerTyrSerArgValIleAlaGlyThrValPheVal---AlaTyrHisArgAspSerTyr 206
QY      484 GAA---AAACAAGTACTTCTTAAACCATCAAAATTCACACGCGCGCAATCTAGGATGG 540
Db      207 GluAsnLysThrMetGlnLeuMetProAspAspTyrSerAsnThrHisSerThrArgTyr 226
QY      541 CACACGACTAATGAGACGTATACCGTGTGGGGATACCATCGATATATATGATATATGATATTTT 600
Db      227 ValThrValLysAspGlnTyrHisSerArgGlySerThrThrLeuTyrArgGluThrCys 246
QY      601 TCGTCAATTGTATAGTAGGAATAATGGATGCGCGCTGTGTTCGTTTCGTTATTCATTTT 660
Db      247 AsnLeuAsnCysMetValThrIleThrAlaArgSerLysTyrProTyrHisPhePhe 266
QY      661 GCAATGGCCAATGGCGACATCGCGAACATATCTCCATTTTATGTTTATGTTTATGTTTATGTTT 720
Db      267 AlaThrSerThrGlyAspValValAspIleSerProPheTyrAsnGlyThrAsnArgAsn 286
QY      721 GCTGCGCGAGAACCCATGGATATCCCGAGGATAATTCACAGGATAATTCAAACAACATAGATACTATTT 780
Db      287 Ala-----SerTyrPhe 290
QY      781 TCAATGGATTGGACAAG----- 798

```

```
Db 291 GlyGluAsnAlaAspLysPhePheLeuProAsnTyrThrIleValSerAspPheGly 310
QY 799 CGTGAAGAACGAGCTTCCAGCTCAAGCGT-----AACTTTCCTCATCACATCACATTC 852
Db 311 ArgProAsnSerAlaLeuGluThrHisArgLeuValAlaPheLeuGluArgAlaAspSer 330
QY 853 ACAGTTGGTGGAGCTGGCTCCAAAGAACTACTCTGTATGTTCAATGACTAAGTGGAAA 912
Db 331 ValIleSerTrpAspIleGlnAspGluLysAsnValThrCysGlnLeuThrPheTrpGlu 350
QY 913 GAGGTGACTGAATGTTGGCTGCAACAGCTAATGGAGATACAGATTTATGGCCCGTGA 972
Db 351 AlaSerGluArgThrIleArgSerGluAlaGluAspSerTyrHisPheSerSerAlaLys 370
QY 973 CTTTCGGCAACGTTTATCAGTAATACGAGTGTGTTGATCCAAATCGCATCATATATAGGA 1032
Db 371 MetThrAlaThrPheLeuSerLysLysGlnGluValAsnMetSerAspSerAlaLeu--- 389
QY 1033 CAATGTATTAACGGCGGACAGCAAGCAGCATCGAGATATTTAGGACAAATATAT 1092
Db 390 AspCysValArgAspGluAlaIleAsnLysLeuGlnIlePheAsnThrSerTyrAsn 409
QY 1093 GACAGTCAGCTCAAGTGGACATGTACAAATATTTCTGGCTCTCGGGGATTTATTGTA 1152
Db 410 GlnThrTyrGluLysTyrGlyAsnValSerValPheGluThrThrGlyGlyLeuValVal 429
QY 1153 GCATATCAGCTGTTCTATCCAAATCCCTGGCTCATATGTACCTCAGAGAATTTGATGAGA 1212
Db 430 PheTrpGlnGlyIleLysGlnLysSerLeuVal-----GluLeuGluArg 444
QY 1213 ---GACACAGCAGCGATGAGATGCTGACCTGGTAAACAAT-----AAGCATGCAATT 1263
Db 445 LeuAlaAsnArgSerSer-----LeuAsnLeuThrHisAsnArgThrLysArgSerThr 462
QY 1264 TATAAGAAAATGCTACCTCATTTGTCACGATTCGCGCGAGATATTCGAATCCACCAAT 1323
Db 463 AspGlyAsnAsnAlaThrHisLeuSerAsnMet----- 473
QY 1324 AGAAAATAACATTAGACGACACACAGCATTAATAATCGACATCTCTGTTCAATTCGCC 1383
Db 474 -----GluSerValHisAsnLeuValTyrAla 482
QY 1384 ATGCTCCCAATTTTATCATATATACAAACCCATATTAATGATATGTTTAGTAGGATT 1443
Db 483 GlnLeuGlnPheThrTyrAspThrLeuArgGlyTyrIleAsnArgAlaLeuAlaGlnIle 502
QY 1444 GCCACAGCTGGTGGCAATTCAGAAATGCAAGTGTGTTTATGGCAGCAAGGATAAAG 1503
Db 503 AlaGluAlaTrpCysValAspGlnArgArgThrLeuGluValPheLysGluLeuSerLys 522
QY 1504 ATTAATCCTAGCGCTACAGCGAGTGCACATTAGGAAGGAGAGTGGCTCCAAAGATGTTG 1563
Db 523 IleAsnProSerAlaIleLeuSerAlaIleTyrAsnLysProIleAlaAlaArgPheMet 542
QY 1564 GGGGATGCTCGCTGCTATCGAGCTGCACTGTATAGATCGGAATCCGCTACCTTGC 1623
Db 543 GlyAspValLeuGlyLeuAlaSerCysValThrIleAsnGlnThrSerValLysValLeu 562
QY 1624 AATTCTATCGGAGTTATCATCCACTAATACATATATATAGCGGACATTTGTTCTATTT 1683
Db 563 ArgAspMetAsnValLysGluSerProGlyArgCysTyrSerArgProValValIlePhe 582
QY 1684 TCATATGGAGAAACCAAGAAACATACAGGACACTCGGTGAAACACAGATGTC 1743
Db 583 AsnPheAlaAsnSerSerTyrValGlnTyrGlyGlnLeuGlyGluAspAsnGluIleLeu 602
QY 1744 CCAACGCTAGAGGCTGTAGAGCGATGCTCGGCTAATATCATCGTAGATATTTCTGTGGA 1803
Db 603 LeuGlyAsnHisArgThrGluGluCysGlnLeuProSerLeuLysIlePheIleAlaGly 622
QY 1804 TCCGGTTATGCTTTATTTGAAAACATAATTTTGTAAAGATGCTAGACGCTCCCGATATA 1863
::: ::::: |||||
```

RESULT 11

US-10-055-364-45

; Sequence 45, Application US/10055364

; Publication No. US20020155433A1

; GENERAL INFORMATION:

; APPLICANT: Patience, Clive

; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use

; FILE REFERENCE: 61750-379

; CURRENT APPLICATION NUMBER: US/10/055,364

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: US/09/612,204

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US/60/142,736

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: US/60/168,532

; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 45

; LENGTH: 830
 ; TYPE: PRT
 ; ORGANISM: Alcelaphine herpesvirus
 US-10-055-364-45

Alignment Scores:

Pred. No.: 3,71e-51 Length: 830
 Score: 758.50 Matches: 224
 Percent Similarity: 43.44% Conservative: 127
 Best Local Similarity: 27.72% Mismatches: 356
 Query Match: 13.39% Indels: 101
 DB: 14 Gaps: 21

US-09-147-052-3 (1-3261) x US-10-055-364-45 (1-830)

```

QY 22 TGCATATTTCTTATAGTATTCTATATGGT-----ACGAACCTCA 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 CysAlaPheLeuLeuLeuPheAlaValLeuLysAsnValPheCysGlnThrProThrSerSer 28
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 64 TCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTGTCTGAG 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 SerGluValGluAspValIleProGluAlaAsnThrValSerAspAsnIleIleArgGln 48
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 124 GAAGAGTCTAGC-----TTTTATCTT 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 GlnArgAsnAsnThrAlaLysGlyIleHisSerAspProSerAlaPheProPheArgVal 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 145 TGTCCCCCACCAGTGGGTTCAACCGTGATCCGCTAGAACCGCGCGAAATGTCGCGAA 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 CysSerAlaSerAsnIleGlyAspIlePheArgPheGlnThrSerHisSerCysProAsn 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 205 CCTAGAAAACCCAGTGGGTGAAGAAATCCGATATATTATTAAGAGAATATACGT 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 89 ---ThrLysAspLysGluHisAsnGluGlyIleLeuLeuIlePheLysGluAsnIleVal 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 265 CCATATAAATTAAGTACGCGCTTATTATAAAATATCATTCAGACGACGATGACG 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 ProThrValPheLysValArgLysThrArgLysIleValThrThrSerThrIleThrAsn 127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 325 GGGACGACATATAGACAGTCACTAATCGATATACAGATAGGACCGCGCTTCCATGAA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 GlyIleThrAlaAspAlaValThrAsnGlnHisValPheSerLysSerValProIleThr 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 385 GAGATCAGGATCTATCGAGCGGAAGGAGATGCTCATCTAAAGCAAGATACCTTGA 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 Glu---ThrArgArgMetAspThrIleThrGlnCysThrAsnSerLeuAspValThrVal 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 445 AACAAATGATATGTTGAAGGTTTGACAGGATGCGGGAGAAAACAAGTACTTCTAAA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 GlyGlyAsnLeuLeuValThrThrAspAsnAspGlySerAsnMetThrValAspLeuGln 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 505 CCATCAAAATTCACACGCCCGAATCTAGGCGCATGGCACACGACTAATGAGAGGTATACC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 ProValAspGlyLeuSerAsnSerValArgArgThrHisSerGlnProGluIleHisAla 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 565 GTGTGGGGATCACCATGGATA-----TATCGAACGGGACCTCCGCTCAATGTATA 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 GluProGly-----TrpLeuLeuGlyGlyThrArgArgThrThrValAsnCysGlu 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 616 GTAGAGAAATGATGCCGCTGTGTTCCGTATTTCATATTTGCAATGGCCCAATGGC 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 ValThrGluThrAspAlaArgAlaValProProPheArgThrPheIleThrAsnIleGly 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 676 GACATCCGCAACATATCTCATTTTATGTTCTATCCACGAGCGTCCGCGAGAACCC 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 AspThrIleGluMetSerProPheThrSerLysAlaTrpAsnGlu-----ThrGluPhe 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 736 ATGGGATATCCCGAGGAATATTCACAACTAGATAGCTATTTTCAATGGATTTGGAC 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 SerGlyGluProAspArgThrLeuThrValAlaLysAspThrArgValValAspThrLys 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 796 AAGCGTCGAAAAGCAACCCCTCCAGTCAACGCGTAACCTTCTCATCATCATCAGCTTCA 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

Db 283 PheArgGlyThrGlnProGlnGlyHisThrArgIlePheValAspLysGluGluThr 302
QY 856 GTTGGGTGGGACCTGCCAAAACTACTCGGTATGTTCAATGACTAGTGAAGAG 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 LeuSerTrpAlaGlnGlnPheArgAsnIleSerThrCysArgTrpAlaHisTrpLysSer 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 916 GTGACTGAATGTTGCGTGCACAGTAAATGGGAGATACAGATTTATGGCCCGTGAACCT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 PheAspAsnAlaIleLysThrGluHisGlyLysSerLeuHisPheValAlaAsnAspIle 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 976 TCGGCAACGTTTATCAGTAATACAGTACGAGTGTGATCCAAATCGCATCATATTAGACAA 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 ThrAlaSerPheThrProAsnThrGln-----ThrArgGluValLeuGlyLys 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1036 -----TGTATTAACCGGAGCAGCAGCAATCGAGCAGATATTTAGCAAAATAT 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 HisValCysLeuAsnAsnThrIleGluSerGluLeuLysSerArgLeu---AlaLysVal 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1090 AATGACAGTCACGTCAGGTTGGACATGTACAATATTCTTGGCTCTCGGGGATTTATT 1149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 AsnAspThrHisSerProAsnGlyThrAlaGlnThrThrLeuThrAsnGlyGlyLeu 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1150 GTAGCATATCAGCTGTTCTATCCAAATCCCTGGCTCATATGTACCTCAGCAATTCATG 1209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 LeuValTrpGlnProLeuValGlnGlnLysLeuAsp----- 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1210 AGACAAACAGGACCGATGAGATGCTCGACCTGTAAACAATAAGCATGCAATTTATAAG 1269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 -----AlaLysGlyLeuLeuAspAlaValLysLys-----Gln 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1270 AAAAATGCTACCTCATTTGTCACGATTCGCGGAGATATTCGAAATGCACCAATAGAAA 1329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GlnAsnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1330 ATAACATACAGCACACACAGCATATTAATCGACATCGTCTGTTCAATTCGCCATGCTC 1389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 SerGlyIleAspValThrThrAlaGluSerThr-----IleLeuLeuThrGlnIle 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1390 CAATTTCTTTATGATCATATACAAACCCATTAATGATATGTTTAGTAGATGCCACA 1449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 GlnPheAlaThrAspThrLeuArgAlaGlnIleAsnAsnValLeuGluLeuSerArg 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1450 GCTTGTGCGAATTCGAAATAGAGAACTGTTTTATGGCAGCAAGGATAAAGATTAAT 1509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AlaTrpCysArgGluGlnHisArgAlaSerLeuMetTrpAsnGluLeuSerLysIleAsn 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1510 CCTAGCCTACAGCAGTGCACCAATAGGAAGAGAGAGTGCCTGCAAGAGATGTTGGGGAT 1569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 ProThrSerValMetSerIleThrGlyArgProValSerAlaLysArgIleGlyAsp 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1570 GTCGCTGCTGTATCGAGCTGCTATAGATCGGAAATCCGTCACCTTTGCAAAATCT 1629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 ValIleSerValSerHisCysValValValValValValValValValValValValVal 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1630 ATCGGAGTT-----ATCACATCCATTAATACATGTTATAGCCGACCATTTGTTTATTT 1683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 MetArgValProGlyArgAspLysThrHisGluCysThrSerArgProProValThrPhe 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1684 TCATATGGAACAAACCAAGCAACATACAGGCAACTCGTGAACCAACAGGATGCTCT 1743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 LysPheIleAsnAspSerHisLeuThrLysGlyGlnLeuGlyValAsnGlnIleLeu 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1744 CCAACGCTAGAGCGCTGAGAGCCATGCTCGCTAAATCATCATGATATTTCTGTTTGGGA 1803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 LeuThrThrAlaValGluIleCysHisGluAsnThrGluHisThrPheGlnGlyGly 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1804 TCCGGTTATGCTTTTATGAAACTATAATTTTGTTAAGATGGTACCGCTCGCGATATA 1863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 AsnAsnMetThrPheThrLysAsnThrArgHisValLysThrMetProValGlyAspVal 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1864 CAGATTCTAGCATTTGTCGAGCTTAATCTAACCTGCTAGAGATCGGAAATTTTG 1923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 AlaThrLeuAspThrPheMetValLeuAsnLeuThrLeuValGluAsnIleAspPheGln 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Qy	94	GAAGTGTGTTCCAGCGTCCTCGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCTTTGTCCTCCCA	151
Db	71	GlUGlnAsnLysAsnIleTyrrGlySerProSerThrPheProTyrrArgValCysSerAla	90
Qy	154	CCAGTGGGTTCACCGTGATCCGCTAGAACCGCGCGAAATGTCGCCGAACCTAGAAA	213
Db	91	SerGlyValGlyAspValPheArgPheGlnThrAspHisValCysProAspAlaSerAsp	110
Qy	214	GCCACCGAGTGGGTGAGGAATCCGATATATTAAAGAGAATATCATGTCATATAAA	273
Db	111	MetValHis---SerGlucLylleuLeuIleTyrrLysGlnAsnIlelleProPheMet	129
Qy	274	TTTAAAGTGCAGCTTTATTATAAAATATCATTCAGACGACGACATGGACGGGACGACA	333
Db	130	PheArgValAlaGlyTyrrArgLysValValThrThrSerThrValTyrrAsnGlyIleTyrr	149
Qy	334	TATAGACAGATCACTAATCGATATACAGATAGGACCGCGTTTCCATTGAGAGATACCG	393
Db	150	SerAspSerIleThrAsnGlnHisThrPheTyrrLysSerIleGluProTrpGlu---Thr	168
Qy	394	GATCTTAATCGACGGCAAGGAGATGCTCATCTAAAGCAAGATACCTTTAGAAAACATGTA	453
Db	169	GlulysMetAspThrIleTyrrGlnCysPheAsnSerLeuArgLeuAsnThrGlyGlyAsn	188
Qy	454	TATGTTGAAGCGTTTGACAGGATGCGGGAGAAAACAAGTACTTCTAAAACCATCAAAA	513
Db	189	LeuLeuThrTyrrValAspArgAspAspIleAsnMetThrValPheLeuGlnProValAsp	208
Qy	514	TTCAACACCGCCGAATCTAGGGCATGGCACAGGACTAATGAGACGCTATACC-----	564
Db	209	GlyValThrProAspValLysArgTyrrGlySerGlnProGluLeuTyrrLeuGluProGly	228
Qy	565	--GTGTGGGGATCACCATGGATATATCGAACGGGAACCTCGTCAATTTGTATATAGAG	621
Db	229	TrpPheTrpGlySer-----TyrrArgArgThrThrValAsnGluLeuMet	245
Qy	622	GAATGGATGCCCGCTGTGTTCCGTATTCTATATTTGCAATGGCCAAATGGCGACATC	681
Db	246	AspMetPheAlaArgSerAsnProPheAspPheValThrAlaThrGlyAspThr	265
Qy	682	CGAACATATCTCCATTTTATGGCTATATCCCCACGAGGCTGCCGCGAAGCCCAATGGGA	741
Db	266	ValGluMetSerProPheTrp-----SerGlyGluAspAspHisGluAsnLysMethis	283
Qy	742	TATCCCCAGGATAATTTCAAACAACATACATAGCTATTTTCAATCGATTTGGACAAGCGT	801
Db	284	GlulysProTrpPheValSerValIleAsnAsnTyrrLysValValAspTyrrGlnAsnArg	303
Qy	802	CGAAAAGCAAGCCTTCCAGGTCAAGGTAACTTTCTCATCATCATCATCTACAGTGTGG	861
Db	304	GlyThrValProLeuGlyLysThrArgIlePheLeuAspArgGluGluTyrrThrLeuSer	323
Qy	862	TGGAGACTGGGCTCCAAAACHTACTCGTGATGTTCATAGCTAACTGGAAGAGGTGACT	921
Db	324	TrpGluLysHisLeuLysAsnMetSerTyrrCysProLeuThrLeuTrpLysAlaPheTyrr	343
Qy	922	GAATGTGTCGTGCAACAGTTAAATGGGAGATACAGATTTATGCGCCGTCGAATTCGGCA	981
Db	344	AsnGlyIleGlnThrGluHisSerGlySerTyrrHisPheValAlaAsnAspIleThrAla	363
Qy	982	ACGTTTATACGT-----AATACGACTGAGTTTGTATCCAAAT	1011
Db	364	SerPheThrThrSerLysGluAspMetLysGluPheAsnThrThrTyrr-----	379
Qy	1018	CGCATCATATTAGGCAATGTATTAAACGCGAGGCAAGAACGAAATCGACGACATATTT	1077
Db	380	-----HisCysLeuAsnGluGluIleLysAlaGluIleGluLysLysTyrr	394
Qy	1078	AGGACAAAATATATGACAGTCACTGACAGTTGGACATGTACATATTTCTTGCTCTC	1131
Db	395	---AlaLysValAsnSerThrHisSerLysTyrrGlyAspLeuLysTyrrPheLysThrAsp	413
Qy	1138	GGGGGATTATTGTAGCATATACAGCCTGTTCTATCCAAATCCCTGGCTCATATGACCTC	1191

US-10-055-364-24
; Sequence 24, Application US/10055364
; Publication No. US20020155433A1

FILE REFERENCE: 61750-379

CURRENT FILING DATE: 2002-01-23

; PRIOR FILING DATE: 2000-07-07

: PRIOR FILING DATE: 1999-07-08
 : PRIOR APPLICATION NUMBER: US/00/142,130

; PRIOR APPLICATION NUMBER: US/60/168,532
: PRIOR FILING DATE: 1999-12-02

```

; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 24
: LENGTH: 965

```

```

; TYPE: PRT
; OPCNTSM: Artificially Generated

```

; FEATURE:

US-10-055-364-24

Alignment Scores:

Score:	754.00	Match
--------	--------	-------

Best Local Similarity: 25.86%

DB: 14 Gaps: 14

US-09-147-052-3 (1-3261) x US-10-055-364-24

QV 49 TATGGTACGAACATCATCTCCG-----

Db
51 TvrGlyThrHisAspSerSerHisGlyGlu

Db 752 Gly-----MetLeuMetIleIleValValIleGlyValLeu 763
 Qy 2179 ATGACATTAGCTTCACTACAAGACTATGCCAAGATT-----GAAGCTAGCTTTATCATCT 2232
 Db 764 PheAlaIleIleThrPheLeuThrLysLysThrLysIleIleThrGluThrAlaProIleLysMet 783
 Qy 2233 GCTTATAGTGAAGCTGAACACAGTTAAACAATAACCTTAAT-----GCAACATTA 2280
 Db 784 IleIleProGluIleAspLysLysLysGluArgGluGlyLysSerGluIleAlaProIle 803
 Qy 2281 GACAACTAAATGGCTAAACATAATTTAGATCAGCCATCAACCAAGCTAATACGGAT 2340
 Db 804 SerGluGluGluLeuGluArgIleValLeuAlaMetHisIleHisGlnAsnSerHis 823
 Qy 2341 AAAACGACTTTTGATATATGAACACCCCAAAATTTAGTGAACCATACAAGCAGCTAAAAACC 2400
 Db 824 MetGluThrLysThrArgLysAspPro-----LysAsp 834
 Qy 2401 ACTTTAGAACAACGCTGCTACTTAACCTTTGAAGTTTGTCACTCAACTGCTTATAATCAAAAT 2460
 Db 835 SerIleLeuThrArgAlaGlnAsnMetLeuArgLysArgSer---GlyTyrSerAsnLeu 853
 Qy 2461 CGCAAT---AATTAGTGGATCTATACAAT 2487
 Db 854 LysAsnAlaGluSerValGluMetLeuAsn 863
 RESULT 13
 US-10-055-364-55
 ; Sequence 55, Application US/10055364
 ; Publication No. US2002015433A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Patience, Clive
 ; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
 ; FILE REFERENCE: 61750-379
 ; CURRENT APPLICATION NUMBER: US/10/055,364
 ; CURRENT FILING DATE: 2002-01-23
 ; PRIOR APPLICATION NUMBER: US/09/612,204
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: US/60/142,736
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: US/60/168,532
 ; PRIOR FILING DATE: 1999-12-02
 ; NUMBER OF SEQ ID NOS: 55
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 55
 ; LENGTH: 793
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Portion of Acelaphine herpesvirus.
 US-10-055-364-55
 Alignment Scores:
 Pred. No.: 1.71e-50 Length: 793
 Score: 750.00 Matches: 229
 Percent Similarity: 44.33% Conservative: 131
 Best Local Similarity: 28.20% Mismatches: 362
 Query Match: 13.24% Indels: 90
 DB: 14 Gaps: 23
 US-09-147-052-3 (1-3261) x US-10-055-364-55 (1-793)
 Qy 136 TTTTATCTTTTGTCCCCACCAGTGGGTTCAACCGTGATCGTCTAGAACCGCGCGAAAA 195
 Db 12 PheArgValCysSerAlaSerAsnIleGlyAspIlePheArgPheGlnThrSerHisSer 31
 Qy 196 TGTCCCGAACCTAGAAAGCCACCGAGTCGGGGTGAAGGATTCGCGATATATTATTAAGAG 255
 Db 32 CysProAsn---ThrLysAspLysGluHisAsnGluGlyIleLeuLeuIlePheLysGlu 50
 Qy 256 AATATCAGTCCTATATAAATTTAAAGTCAGCCTTTATATATAAAATATCATTCAGACGACG 315
 Db 51 AsnIleValProIleValPheLysValArgLysIleValThrSerThr 70

Db 758 GluGluGlnGlySerThrSerGlyProAlaAspLeuAsnLysAlaLysAsnValLeu 777
QY 2377 GAA-----GCATACAAAGCAGCTAAACCACT 2403
Db 778 ArgArgAlaGlyTyrLysProLeuLysArgThr 789

RESULT 14

US-10-055-364-40
; Sequence 40, Application US/10053364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Bovine herpesvirus 4
US-10-055-364-40

Alignment Scores:

Pred. No.:	1,41e-49	Length:	844
Score:	738.50	Matches:	217
Percent Similarity:	41.98%	Conservative:	126
Best Local Similarity:	26.56%	Mismatches:	337
Query Match:	13.04%	Indels:	137
DB:	14	Gaps:	22

US-09-147-052-3 (1-3261) x US-10-055-364-40 (1-844)

QY 55 ACGAACTCTCCGAGTACCCAAATGTG---ACATCAAGAGAAGTGTTCGAGCGTC 111
Db 62 SerThrSerLysProSerThrAspAsnGlnGlyThrSerThrProThrIleProThrVal 81
QY 112 CAGTTGCTGAGGAAGAGTCTACGTTTAT-----CTTTGCCCCACCGATGGGT 162
Db 82 ThrAspAspThrAlaSerLysAsnPhetYrLysTyrArgValCysSerAlaSerSer 101
QY 163 TCA---ACCGTGATCCGTCTAGAACCGCGCGGAAAATGTCGGAACCTAGAAAGCCACC 219
Db 102 SerGlyGluLeuPheArgPheAspLeuAspGlnThrCysProAsp---ThrLysAspLys 120
QY 220 GAGTGGGTGAAGGAATCCGATATATTATAAGAGAATATCAGTCCCATATAAATTTAA 279
Db 121 LysHisValGluGlyIleLeuLeuValLeuLysAsnIleValProTyrIlePheLys 140
QY 280 GTGAGCTTTATTATAAAATATCATTCAGACGACATGACGCGGACGACATATACA 339
Db 141 ValArgLysTyrArgLysIleAlaThrSerValThrValTyrArgGlyTyrSerGlnAla 160
QY 340 CAGATCATAATCGATATACAGTAGGACCGCCGTTTCCATGAAGACATCAGGATCTA 399
Db 161 AlaValThrAsnArgAspIleSerArgAlaIleProTyrAsnGluIleSer---Met 179
QY 400 ATCGACGCGAAAGGAAGATGCTCATTAAGCAAGATACCTTAGAAACAATGTATGTT 459
Db 180 IleAspArgThrTyrHisCysPheSerAlaMetAlaThrValIleAsnGlyIleLeuAsn 199
QY 460 GAAGCGTTTGACAGGATCGCGGAGAAACAAAGTACTTCTAAACCACTCAAAATTCAC 519
Db 200 ThrTyrIleAspArgAspSerGluAsnLysSerValProLeuGlnProValAlaGlyLeu 219

QY 520 ACGCCCGAATCTAGGCGATGGCCACAGCACTAATGAGAGCTATACCGTGTGGGA---TCA 576
Db 220 ThrGluAsnIleAsnArgTyrPheSerGlnProLeuIleTyrAlaGluProGlyTyrPhe 239
QY 577 CCATGGATATATCGAACGGAACTCCGTCATTTATAGTAGAGCAATGATGATCCCGC 636
Db 240 ProGlyIleTyrArgValArgThrThrValAsnCysGluValValAspMetTyrAlaArg 259
QY 637 TCTGTGTTTCGCTATTTTCATATTTTCCAAATGGCCATGGCCATCGGCAACATATCTCA 696
Db 260 SerValGluProTyrThrHisPheIleThrAlaLeuGlyAspThrIleGluIleSerPro 279
QY 697 TTT-----TATGGTCTATCCCCACAGAGGCTCCCGAGAA 732
Db 280 PheCysHisAsnAsnSerGlnCysThrThrGlyAsnSerThrSerArgAspAlaThrLys 299
QY 733 CCCATGGGATATCCCGAGATAATTTCAACAACTAGATAGTAGTATTTTCAATGGATTG 792
Db 300 Val-----TrrIleGluGluAsnHisGlnThrValAspTyr----- 311
QY 793 GACAAGCTCGAAAGCAAGCCTCCAGTCAAGCGTAACCTTTCTCATCACATCACACTTC 852
Db 312 ---GluArgArgGlyHisProThrLysAspLysArgIlePheLeuLysAspGluGluTyr 330
QY 853 ACAGTTGGTGGGACTGGCTCCAAAACACTACTCGTGTATGTTCAATGACTAAGTGGAA 912
Db 331 ThrIleSerTrrLysAlaGluAspArgGluArgAlaIleCysAspPheValIleTrrLys 350
QY 913 GAGGTGACTGAAATGTTGCTGCAACAGTTAATGGAGATACAGATATTTATGCGCGTGA 972
Db 351 ThrPheProArgAlaIleGlnThrIleHisAsnGluSerPheHisPheValAlaAsnGlu 370
QY 973 CTTTCGGCAACGTTTATCAGTAATACG-----ACTGAGTTTCATCCAAATCGCATC 1023
Db 371 ValThrAlaSerPheLeuThrSerAsnGlnGluGluThrGluLeuArgGlyAsnThrGlu 390
QY 1024 ATATTAGGACCAATGTATTAACGCGCAGCAGCAGCAATCGACAGCATATTTAGGACA 1083
Db 391 IleLeu---AsnCysMetAsnSerThrIleAsnGluThrLeuGluGluThrValLys--- 408
QY 1084 AAATAATAATGACAGTCAGTCAGTCAAGTTGGACATGTACAATATTTCTGCTCGGGGA 1143
Db 409 LysPheAsnLysSerHisIleArgAspGlyGluValLysTyrTyrLysThrAsnGlyGly 428
QY 1144 TTTATTGACATATCAGCTGTTTCTATCCAAATCCCTGGCTCATATGTAACCTACAGAA 1203
Db 429 LeuPheLeuIleTrrPglAlaMetLysProLeuAsnLeuSerGluHisThrAsnTyrThr 448
QY 1204 TTGATGAGACACAACAGCAGCAGATGATGCTCGACCTGGTAAACAATAAGCATGCAAT 1263
Db 449 IleGluArgAsnAsnLysThr-----GlyAsnLys----- 458
QY 1264 TATAAGAAAAATGCTACCTCATTTGACGATTCGCGGAGATATTCGAAATGCACCAAT 1323
Db 459 -----SerArgGlnLysArgSerVal----- 465
QY 1324 AGAAAAATAACATATAGACACACAGCTATTAATCGACATGCTGTTTCAATTCGCC 1383
Db 466 -----AspThrLysThrPheGlnGlyAlaLysGlyLeuSerThrAla 479
QY 1384 ATGCTCCAAATTTCTTATGATCATATACAAACCCATATTAATGATATGTTTAGTAGATT 1443
Db 480 GlnValGlnTyrAlaTyrAspHisLeuArgThrSerMetAsnHisIleLeuGluGluLeu 499
QY 1444 GCCACAGCTTGGTGGCAATTCGAGAAATAGAACCTTTTATGCAACAGGATAAAG 1503
Db 500 ThrLysThrTrrPcyArgGluGlnLysLysAspAsnLeuMetTrrTrrGluLeuSerLys 519
QY 1504 ATTAATCTAGCGTACAGCGAGTCCAACTATTAGGAGAGAGTGGCTGCAAAAGATTG 1563
Db 520 IleAsnProValSerValMetAlaAlaIleTyrGlyLysProValAlaValLysAlaMet 539
QY 1564 GGGGATGTCGTGTCGAGCTGCTATGATGCGGAATCCGCTCACTTGGCA 1623


```

Db 540 GlyAspAlaPheMetValSerGluCysIleAsnValAspClnAlaSerValAsnIleHis 559
QY 1624 AATTCTATCGAGTTATCATCATCCACTAATATACATATATAGCGGACCATTTGTTCTATTT 1683
Db 560 LysSerMetArgThr---AspAspProLysValCysTyrSerArgProLeuValThrPhe 578
QY 1684 TCATATGAGAGAAACCAAGAACATACAGGAGCAACTCGGTGAAACACAGATGCTTT 1743
Db 579 LysPheValAsnSerThrAlaThrPheArgGlyGlnLeuGlyThrArgAsnGluIleLeu 598
QY 1744 CCAACCTAGAGGCTGTAGAGCATCTCGCTAATCATCGTAGATATTTCTGTTTGA 1803
Db 599 LeuThrAsnThrHisValGluThrCysArgProThrAlaAspHisTyrPhePheValLys 618
QY 1804 TCCGGTATGCTTTATTTGAAACATAATTTTGTAGATGCTGACGCTGCGGATATA 1863
Db 619 AsnMetThrHisTyrPheLysAspTyrLysPheValLysThrMetAspThrAsnAsnIle 638
QY 1864 CAGATTGCTAGACATTTCTGAGCTTAATCTAACCTGCTAGAGATCGGGAATTTTG 1923
Db 639 SerThrLeuAspThrPheLeuThrLeuAsnLeuThrPheIleAspAsnIleAspPheLys 658
QY 1924 CCTTATCCGTTTACACAAAGAGTTCGGTGTGTTGGTCTATTGGATTATGCAGAA 1983
Db 659 ThrValGluLeuThrSerGluThrGluArgLysMetAlaSerAlaLeuAspLeuGluThr 678
QY 1984 GTAGCTCGCGCAATCACTACATGAACCTTAATTTTATGACATAACAAAGTAATAGAA 2043
Db 679 MetPheArgGluTyrAsnTyrTyrThrGlnLysLeuAlaSerLeuArgGluAspLeuAsp 698
QY 2044 -----GTGGATACAAATTAACCGGGGCTGCAGGAATTCGGCTGTATCTATTACTAAA 2097
Db 699 AsnThrIleAspLeuAsn----- 704
QY 2098 AAAGATGCAACCAACCAATAATAGCCCAACCAATTAAGACCGCGAATGGATTAACA 2157
Db 705 ArgAspArg----- 707
QY 2158 GATCTAATCAATGCTAAAGCATGACATTAGCTTCACTACAGACTATGCCAAGATTGAA 2217
Db 708 ---LeuValLysAspLeuSerGluMetMetAlaAspLeuGlyAspIleGlyLysVal--- 725
QY 2218 GCTAGTTTATCATCTCTTATAGTGAAGCTGAACAGT----- 2256
Db 726 -----ValValAsnThrPheSerGlyIleValThrValPheGlySerIleValGlyGly 743
QY 2256 ----- 2256
Db 744 PheValSerPhePheThrAsnProIleGlyValThrIleIleLeuLeuIleVal 763
QY 2257 -----AACAAATAACCTTAATGCAACATTAGAA 2283
Db 764 ValValPheValValPheIleValSerArgArgThrAsnAsnMetAsnGluAla----- 781
QY 2284 CAACTAAAATGGCTAAAATCTAATTTAGATAGCCATCAGCAACCAACTAAT 2334
Db 782 ProIleLysMetIleTyrProAsnIleAspLysAlaSerGluGlnGluAsn 798

```

RESULT 15

```

US-10-055-364-38
; Sequence 38, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patient, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08

```

```

; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Rhesus monkey rhadinovirus
US-10-055-364-38

```

Alignment Scores:

```

Pred. No.: 1,47e-48 Length: 808
Score: 725.50 Matches: 208
Percent Similarity: 45.37% Conservative: 115
Best Local Similarity: 29.21% Mismatches: 306
Query Match: 12.81% Indels: 83
DB: 14 Gaps: 19

US-09-147-052-3 (1-3261) x US-10-055-364-38 (1-808)
QY 13 AGCGGAATTCATATTTTCCCTATATAGTATTCTATATGTTAGCAACTCATCTCCGAGT 72
Db 8 ArgArgLeuLeuArgAlaTrpValIleIleAlaIleGlyThrAlaValGly----- 25
QY 73 ACCCAAAATGTGACATCA---AGAGAAGTTGTTTCGAGCGTCCTGCTGAGGAAGAG 129
Db 26 ---GluAsnValThrThrProLysGlyAlaThrThrAlaLysProThrProGlyPro 44
QY 130 TCTACG-----TTTATCTTTGT 147
Db 45 SerThrProThrProGluAsnProProArgAlaGluAlaPheLysPheArgValCys 64
QY 148 CCCCCACCATGTTCAACCGTCCGCTAGACCGCGCGGAAATATCCCGAACCT 207
Db 65 SerAlaSerAlaThrGlyGluLeuPheArgPheAsnLeuLysThrCysProGlyThr 84
QY 208 AGAAAAGCCACCGAGTGGGTGAAGATTCGCGATATATTATTAAGAGATATCAGTCCA 267
Db 85 GluAspLysThrHis---GlnGluGlyIleLeuMetValPheLysLysAsnIleValPro 103
QY 268 TATAAATTAAGTGACGCTTTTATTAATAAATATCATTCAGACGACGACATCGAGGG 327
Db 104 HisIlePheLysValArgTyrArgLysValAlaThrSerValThrValTyrArgGly 123
QY 328 ACACATATAGACAGATCACTAATCATATACATAGATAGGACGCGCTTCATTTGAAGAG 387
Db 124 TrpThrGluThrAlaValThrGlyLysGlnGluValIleArgProValProGlnTyrGlu 143
QY 388 ATCAGCGATCTAATTCAGCGCAAGGAAGATGCTCATCTAAACAAAGATACCTTAGAAC 447
Db 144 IleAsnHisMet---AspThrThrTyrGlnCysPheSerSerMetArgValAsnValAsn 162
QY 448 AATGTATATGTTGAACGGTTTGACAGGATCGCGGAGAGAAACAAAGTACTTCTAAAACCA 507
Db 163 GlyIleValAsnThrTyrThrAspArgAspPheThrAsnGlnThrValPheLeuGlnPro 182
QY 508 TCAAAATTCACACACGCGCAATCTAGGCGTACACGACACTAATGAGACGATATACCGTG 567
Db 183 ValGluGlyLeuThrAspAsnIleGlnArgTyrPheSerGlnProValLeuTyrThrThr 202
QY 568 TGGGGA---TCACCATGGATATATCAACGCGGAACTCCGCTCAATTTGATAGAGAGAA 624
Db 203 ProGlyTrpPheProGlyIleTyrArgValArgThrThrValAsnCysGluIleValAsp 222
QY 625 ATGGATGCGCGCTCTGTTTCGCTATTATATTTTGCATATGCGCAATGCGCATCGCG 684
Db 223 MetIleAlaArgSerAlaGluProTyrSerTyrPheValThrAlaLeuGlyAspThrVal 242
QY 685 AACATATCTCCATT-----TATGGTCTATCCCGACACGAGCTGCGCGAGAA 732
Db 243 GluValSerProPheCysHisAsnAspSerThrCysSerValAlaGluLysThrGluAsn 262
QY 733 CCCATGGGATATCCCGCAGGATAATTTCAAAACACTAGATAGTATTTTCAATGGATTG 792

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:41:47 ; Search time 51.3424 Seconds
(without alignments)
5374.731 Million cell updates/sec

Title: US-09-147-052-3

Perfect score: 5665

Sequence: 1 atgcacatttttagcgga.....gacccgggtacatttttataa 3261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framet_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09147052/runat_08102003_154342_29880/app_query.fasta_1.5980
-DB=Issued Patents_AA -OFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPGL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HCAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052 -CGN_1_126 -runat_08102003_154342_29880 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3601	63.6	865	1	US-07-803-633A-13
2	1968	34.7	879	1	US-08-220-151-2
3	1968	34.7	879	1	US-08-220-151-3
4	1968	34.7	879	1	US-08-413-118-2
5	1968	34.7	879	1	US-08-413-118-3
6	1968	34.7	879	1	US-08-413-118-106
7	1968	34.7	879	3	US-08-473-446-2
8	1968	34.7	879	3	US-08-473-446-3
9	1968	34.7	879	3	US-08-473-446-106
10	1913	33.8	615	2	US-08-525-742-10
11	1859	32.8	610	2	US-08-525-742-8
12	1846	32.6	1041	1	US-08-220-151-4

13	1846	32.6	1041	1	US-08-413-118-4
14	1846	32.6	1041	3	US-08-473-446-4
15	1779	31.4	913	3	US-09-232-468A-2
16	1779	31.4	913	4	US-09-784-984B-49
17	1777.5	31.4	868	1	US-08-220-151-7
18	1777.5	31.4	868	1	US-08-413-118-7
19	1777.5	31.4	868	3	US-08-804-439A-21
20	1777.5	31.4	868	3	US-08-473-446-7
21	1777.5	31.4	868	3	US-08-720-229-21
22	1775	31.3	913	1	US-08-220-151-6
23	1775	31.3	913	1	US-08-413-118-6
24	1775	31.3	913	1	US-08-473-446-6
25	1755	31.0	913	6	5196516-8
26	1748	30.9	933	2	US-08-682-847-2
27	1743	30.8	943	3	US-08-911-321-4
28	1741.5	30.7	980	1	US-08-220-151-5
29	1741.5	30.7	980	1	US-08-413-118-5
30	1741.5	30.7	980	3	US-08-473-446-5
31	1739.5	30.6	891	1	US-08-042-747A-6
32	1736	30.6	904	4	US-09-894-998A-18
33	1735	30.6	904	3	US-08-632-537-2
34	1735	30.6	904	5	PCT-US96-05316-2
35	1735	30.6	904	6	5244792-3
36	1719.5	30.4	885	1	US-08-042-747A-8
37	1714.5	30.3	885	3	US-08-804-439A-23
38	1714.5	30.3	885	3	US-08-720-229-23
39	1705.5	30.1	904	3	US-08-632-537-1
40	1705.5	30.1	904	5	PCT-US96-05316-1
41	1702.5	30.1	903	3	US-08-804-439A-22
42	1702.5	30.1	903	3	US-08-720-229-22
43	1699.5	30.0	903	1	US-08-220-151-8
44	1699.5	30.0	903	1	US-08-413-118-8
45	1699.5	30.0	903	3	US-08-473-446-8

ALIGNMENTS

RESULT 1

US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MARER'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103p
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848

Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 49, Appli
Sequence 7, Appli
Sequence 21, Appli
Sequence 21, Appli
Sequence 6, Appli
Sequence 6, Appli
Patent No. 5196516
Sequence 2, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 18, Appli
Sequence 2, Appli
Sequence 2, Appli
Patent No. 5244792
Sequence 8, Appli
Sequence 23, Appli
Sequence 23, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli

```

; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-803-633A-13

Alignment Scores:
Pred. No.: 1.4e-294 Length: 865
Score: 3601.00 Matches: 719
Percent Similarity: 87.8% Conservative: 16
Best Local Similarity: 85.90% Mismatches: 60
Query Match: 63.57% Indels: 42
DB: 1 Gaps: 8

US-09-147-052-3 (1-3261) x US-07-803-633A-13 (1-865)

QY 1 ATGCACATATTTAGCGGAAATTCATATTTTCCCTATAGTTATTTCTATATGTTAGTACGAC 60
Db 1 MetHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
QY 61 TCATCTCCGAGTACCAAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
Db 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCCAAGTGGGTTCAACCGTGTATCGGTCTA 180
Db 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60
QY 181 GAACCGCCGCAAAATGTCCTCAAGTACGATCAATTAATTTAAAGTCAGCGTTTATTATAAAT 240
Db 61 GluProProArgGlyCysProGluProArgLysAlaThrGluTrpGlyGlyIleAla 80
QY 241 ATATTATTTAAAGAAATATCATCTCAATATAATTTAAAGTCAGCGTTTATTATAAAT 300
Db 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
QY 301 ATCATTCAGCACGACATGAGGGGAGGAGATATAGACAGATCACTAATCGATATACA 360
Db 101 IleIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120
QY 361 GATAGGAGCGCGTTCCTATCAAGAGATCACGGATCTAATCGCGCAAGCAAGATGC 420
Db 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
QY 421 TCATCTAAGCAAGATACCTTGAACAATGTATATGTTGAAGCGTTTGACAGGATCG 480
Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla 160
QY 481 GGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCTAGGSCATGG 540
Db 161 GlyGluLysGlnValIleLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
QY 541 CACAGCACTAATGAGACGTATACCGTGTGGGATCACCATGGATATATCGAACGGGAACC 600
Db 181 HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
QY 601 TCCGTCAATGTATAGTAAAGAAATGGATGCCCTCTGTGTTTCCGTATTATATTTT 660
Db 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
QY 661 GCAATGGCCATGGGACATCGGAACATATCTCCATTTATGGTCTATGCCACACAGAG 720
Db 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu 240
QY 721 GCTGCGCGAGAACCCATGGGATATCCCGAGATATTTCAACAACTAGATAGCTATTTT 780
Db 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe 260
QY 781 TCAATGGATTGGACAAGCGTCGAAAGCAAGCCCTCCAGTCAAGCGTAACCTTCTCATC 840

```

```

Db 261 SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPheLeuIle 280
QY 841 ACATCACACTTCACAGTTGGGTGGGACTGGCTCCAAAACACTACTGCTGTATGTTCAATG 900
Db 281 ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet 300
QY 901 ACTAAGTGGAAAGAGGTGACTGAAATGTTGCGTGCACAGAGTTAATGGGAGATACAGATT 960
Db 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
QY 961 ATGGCCCGTGAACCTTCGCGCAACGTTTATCAGTAATACAGCTAGGTGTTGCTCAATCGC 1020
Db 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
QY 1021 ATCATATTAGGACAAATGTATTAAACGCGAGGAGCAGACAGCAATCGAGCAGATATTAGG 1080
Db 341 IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
QY 1081 ACAAAATATATGACAGCTCACGTCAGGTTGGACATGTACAATATTTCTGGCTCTCGGG 1140
Db 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
QY 1141 GGATTTATTGTAGCATATCAGCCTGTCTCTATCCAAATCCCTGGCTCATATGTACCTCAGA 1200
Db 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
QY 1201 GAATTGATGAGACACAACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATGCA 1260
Db 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
QY 1261 ATTTATAGAAAAATGCTTACCTACCTGTCACGATTGCGGCGAGATATTGCGAATGCACCA 1320
Db 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro 440
QY 1321 AATAGAAAAATAACATTAGACGACACACAGCTATTAAATCGACATCGTCTCTCAATTC 1380
Db 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe 460
QY 1381 GCCATGCTCCAAATTTCTTTATGATCATATACAAACCCATATTAAATGATATGTTTAGTAGG 1440
Db 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
QY 1441 ATTGCCACAGCTTGTGGCAATTCAGATACAGAACTTGTTTATGGCCACCAAGGATA 1500
Db 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle 500
QY 1501 AAGATTAATCTTAGCGCTACAGCGAGTGCACATTAGGAAGGAGAGTGGCTGCTCAAGATG 1560
Db 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
QY 1561 TTGGGGGATGCTGCTGTATCGAGCTGCTGCTATAGATGCGGAATCCGCTCACTTTG 1620
Db 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
QY 1621 CAAAATTTCTATCGAGTTATCATCCACTATACATGCTTATAGCCGACCATTTGTTCTA 1680
Db 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTCATATGGAGAAAACCAAGAAACATACAGGACCACTCGGTGAAAACAACAGGTG 1740
Db 561 PheSerTyrGlyLysGlnGlnGlyAsnIleGlnGlyGlnLeuGlyGlnAsnGlnLeu 580
QY 1741 CTTCAACGCTAGAGGCTGTAGAGCCATGCTCGGCTAATCATCGTAGATATTTTCTGTTT 1800
Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe 600
QY 1801 GGATCCGCTTATGCTTTTATTTCAAACATATAATTTTGTTAAGATGGTAGACCTCCGAT 1860
Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAsp 620
QY 1861 ATACAGATTGTCAGACATTTGTCAGGCTTATCTAACCTCTAGAGATCGGAAATTT 1920
Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGluAspArgGluIle 640

```


Qy	754	AAITTTCAACAACATGATAGCTATTTTTCATTTGGAAGCGTCGAAAGCAAGC	813
Db	267	ArgPheGlnGlnIleGluGlyTyrProIleAspLeuaspThrArgLeuGlnLeuGly	286
Qy	814	CTTCCAGCTCAAGCGTAACCTTCATCACATCCACATTCACAGTTGGTGGGACTGGCGT	873
Db	287	AlaProValSerArgAsnPheLeuThrGlnHisValThrValAlaIatPasnTrpVal	306
Qy	874	CCAAAATACTCTCGTGTATGTTCAATGACTAAAGTGGAAAGAGGTGACTGAAATGTTCCGT	933
Db	307	ProIysIleArgGluValCysThrLeuAlaIysTrpArgGluIleaspGluIleIleArg	326
Qy	934	GCAACAGTTAATGGGAGATCAGATTTATGGCCGTGAACCTTCGGCAACGTTTATCAGT	993
Db	327	AspGluTyrLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer	346
Qy	994	AATACGACTCAGTTGATCCAAATCCGATCATATTAGACAAATGTTAATAACCGGAGCA	1053
Db	347	AspThrThrGlnPheAspIleaspArgValLysLeuSerAspCysAlaLysArgGluAla	366
Qy	1054	GAAGCAGCAATCGACGAGATATTAGACAAAATAATATGACAGTACCGTCACAGTTGGA	1113
Db	367	IleGluAlaIleAspLysIleTyrLysLysLysTyrAsnLysThrHisIleGlnThrGly	386
Qy	1114	CATGTACAATATTTCTGGCTCTCGGGGGATTTATTTAGCATATACAGCTGTTTCTATCC	1173
Db	387	GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer	406
Qy	1174	AAATCCCTGCTCATATGTACCTCAGAGAATGTATGATGAGACACACAGCCGATGACATG	1233
Db	407	AsnGluLeuAlaLysLeuTyrIleasnGluLeuValArgSerAsnArgThrValAspLeu	426
Qy	1234	CTCGACCTGTTAAACAATAGCATGCAATTTATAAGAAAATGCTACTCTATGTTCCACGA	1293
Db	427	LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu	446
Qy	1294	TTGCGGGCAGATATTCAAATGCACCAAAATAGAAAATACATTTAGACGACACACAGCT	1353
Db	447	AsnLysArgSerLysArgAsnIleGluGlyGlyIleGluAsnValAsnAsnSerThrIle	466
Qy	1354	ATTAAATCGACATCGCTGTTCAATTCGCCCATGCTCCAATTTCTTTTATGATCATATACAA	1413
Db	467	IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln	486
Qy	1414	ACCCATATTAAATGATATGTTTAGTAGGATGCCACAGCTTGGTCGCAATTCACAGATAGA	1473
Db	487	SerHisValAsnGluMetLeuSerArgIleAlaThrAlaIatPcysAsnLeuGlnAsnLys	506
Qy	1474	GAACTTGTTTATGGCAGGAGGATAAAAGATTAACTTAGCCGCTACAGGAGTGCACAAC	1533
Db	507	GluArgThrLeuTrpAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla	526
Qy	1534	TTAGGAAGAGAGTGGCTGCAAAAGATTTTCGGGGATGTCGCTGCTATTCGAGCTGCAC	1593
Db	527	MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal	546
Qy	1594	GCTATAGATCGGGAATCCGCTACTTTTGCAAAATCTATGCGAGTTATFCATCCACTAAT	1653
Db	547	AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr	566
Qy	1654	ACATGTTATAGCCGACCATGGTTCTATTTTCATATGCGAGAAAACCAAGGAAC--ATA	1710
Db	567	ThrCysTyrSerArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrIle	586
Qy	1711	CAGGGACAACCTCGGTGAAAACAACAGAGTTGCTTCCAAACGCTAGAGCGCTAGAGCCATGC	1770
Db	587	GluCylGlnLeuGlyGluAsnAsnGluLeuValGluArgLysLeuIleGluProCys	606
Qy	1771	TCGGCTAATCATCGTAGATATTTCTGTTTGGATCCGGTTATGCTTTATTTGAAACTAT	1830
Db	607	ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr	626

```

1931 AATTTTGTAAAGATGGTAGCGCTGCCGATATACAGATTCTGCACAAATTGTTCGAGCTT 1899
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
627 AlaTyrValArgLysValProLeuAsnGlnLeuMetIleSerAlaTyrValAspLeu 646
1991 AATCTAACCCCTGCTAGAAAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAGAGAG 1950
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
647 AsnIleThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu 666
1951 TTGGCTGATGTTGGTGTATTGGATTATGCAGAAAGTAGCTCGCGCAATCAACTACATGAA 2010
      ||| ||| ||| ::::: ||| ::::: ||| ::::: |||
667 LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisAla 686
2011 CTTAAATTTTATGACATAAACAAGTAATAAGAAGTGCATACAAATTACGCGGCGCTGCAG 2070
      ||| ||| ||| ::::: ||| ::::: ||| ||| :::::
687 LeuLysPheTyrAspIleAspSerValValLysValAspAsnValValIleMetArg 706
2071 GAATTCGGCTGTATGCTCTATTACTAAAAAAGATGCAAACCCAAATATATGCCAAACCCAA 2130
      ||| ||| ||| :::::
707 GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal 726
2131 TTAGAAGCAGCGGAATGGAGTTTAACAGATCTAATCAATGCTTAAGCGATGACATAGCT 2190
      ||| ||| ||| ||| ||| ||| :::::
727 LeuGlyAlaAla ::::: --AsnAlaValIleAlaThrValSer 738
2191 TCACACAAGACTATGCTCAAG-- ::::: -----ATTGAA 2217
      ::: |||
739 GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeu 758
2218 GCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACAAATACCTTAATGCAACA 2277
      ||| ||| ::::: ||| :::::
759 AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro 778
2278 TTAGAACAACATAAAATGGCTAAACT--AATTAGAAATCAGCCATCAACCAAGCTAAT 2334
      ::::: ||| ||| ||| ::::: ::::: |||
779 MetLysAlaLeuTyrProValThrThrLysAsnLeuLysGluSerValLysAsnGlyAsn 798
2335 ACGGATAAACGACTTTTGTATAATGACACCCAAATTTAGTTGAA----- 2379
      ::: ||| ::: ||| ::: |||
799 SerGlyAsnAsnSerAspGlyGluGluAsnAspAsnIleAspGluGluLysLeuGln 818
2380 -----GCATCAAGAGCTAAACCACTTTAGAACACGT----- 2415
      ||| ::::: ||| ::::: |||
819 GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnGluHis 838
2416 -----GCTACTACCTTGAAGCTTTGTCTCA 2439
      ||| ::::: ||| ||| |||
839 LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer 858
2440 -----TCACTGCTTATATCAAAATTCGCAAT 2466
      ||| ::::: |||
859 LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn 871

RESULT 3
US-220-151-3
; Sequence 3, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc AND gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```


QY	394	GATCTAATCGAGCGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTTAGAAACAATGTA	453	1474	GAACCTGTTTATGCGCAGCAAGGGATAAAGATTAACTCTAGCGCTACGCGTAGTGCACA	1533
Db	148	GluLeuIleAspArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr	167	507	GluArgThrLeuTrpAsnGluValMetLysLeuAsnProThrSerValalaSerValala	526
QY	454	TATGTTGAAGCGTTTGACAGGATCGCGGAGAGAAACAAGTACTTCTAAACCCATCAAAA	513	1534	TTAGCAAGAGAGTGGCTGCAAGATGTGGGGGATGCGCTGCTGTATCGAGCTGCACT	1593
Db	168	GluPheThrAlaPheAspGluAspProArgGluValHisLysLeuProSerLys	187	527	MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal	546
QY	514	TTCAACAGCCCGAATCTAGGCGATGGCACAGCACTAATGAGACCTATACCGTGGGGA	573	1594	GCTATAGATGCGGAATCCGTCATTTGCAAAAATCTATGCGAGTTATCACATCCACTAT	1653
Db	188	PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly	207	547	AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr	566
QY	574	TCACCATGGATATATCGAACCGGAACCTCCGTCATATGTATAGTAGAGAAATGATGCC	633	1654	ACATGTTATAGCCGACCATTTGCTTCTATTTTTCATATGAGAGAAACAACGAGAAAC	1710
Db	208	GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluValAspAla	227	567	ThrCysTyrSerArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrIle	586
QY	634	CCCTCTGTGTTCCGTATTCATATTTTGCATGGCCAATGGCGACATCGGCNACATATCT	693	1711	CAGGACACACTCGGTGAAACAACACGATGTGCTTCCAAACGCTAGAGGCTCAGAGCATC	1770
Db	228	ArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspIleHisMetSer	247	587	GluGlyGlnLeuGlyGluAsnAsnGluLeuValGluArgLysLeuIleGluProCys	606
QY	694	CCATTTTATGGTCTATCCACACAGAGGCTGCCGAGAACCCATGGGATATCCCCAGGAT	753	1771	TCGCGTAATCATCGTAGATATTTTCTGTTTGGATCCGTTATGCTTTATTTGAAACTAT	1830
Db	248	ProPhePheGlyLeuArg---AspGlyAlaHisThrGluTyrIleSerTyrSerThrAsp	266	607	ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr	626
QY	754	AATTTCAACAACACTAGATAGCTATTTTCAATGGATTTGGACAACGCTCGAAAGCAAGC	813	1831	AATTTTGTAAAGTGGTAGACGCTCGGATATACAGATTTGTCAGACATTTGTCAGCTT	1890
Db	267	ArgPheGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGlnLeuGly	286	627	AlaTyrValArgLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu	646
QY	814	CTTCAGTCAACGCTAACTTCTCATCATCATACACTTCACAGTTGGTGGGACGTTGGCT	873	1891	AATCTAACCTCTAGACATCGGAAATTTTGCCTTTATCGTTTACACAAAAGAGAG	1950
Db	287	AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaTrpAsnTrpVal	306	647	AsnIleThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu	566
QY	874	CAAAAACACTACTCGGTATGTTCAATAGTAAAGTGAAGAGTGCTGACTGAAATGTCGT	933	1951	TTGCGTGATGTGGTGATTGATTGATGAGAAAGTACGCTCGCGCAATCACTACATGAA	2010
Db	307	ProLysIleArgGluValCysThrLeuAlaLysTyrArgGluLeuAspCysIleLeuArg	326	667	LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisAla	686
QY	934	GCAACAGTTAATGGGAGATACAGATTATGCGCGTGAACTTTTCGCAACGTTTATCAT	993	2011	CTTAAATTTTATGACATAAACAAGTAATAGTAGGATACAAATTTACCGGGGCTGAC	2070
Db	327	AspGluTyrLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer	346	687	LeuLysPheTyrAspIleAspSerValValLysValAspAsnAsnValValIleMetArg	706
QY	994	AATAGCACTGATGTTGATCCCAATCGCATCATATTAGGCAATGTATTAAACCGCAGCA	1053	2071	GAATTCGGCTGTATGCTATTACTAAAAAGATGCAACCAATAATGCGCAACCCAA	2130
Db	347	AspThrThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla	366	707	GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal	726
QY	1054	GAAGCAGCAATCGCAGATATTTAGGCAAAATATAATGACAGTCACGCTCAAGGTGGA	1113	2131	TTAGAGCAGCGCGAATGAGTAAACAGATCAATCAATGCTCTAAAGCGATGATAGCT	2190
Db	367	IleGluAlaIleAspLysIleTyrLysLysLysTyrAsnLysThrHisIleGlnThrGly	386	727	LeuGlyAlaAla-----AsnAlaValIleAlaThrValSer	738
QY	1114	CATGTACAATATTTCTGGCTCTCGGGGATTTATGTAGCATATACGCTGTCTATGCC	1173	2191	TCACTACAAGACTATGCCAAG-----ATTGAA	2217
Db	387	GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleAlaPheArgProMetIleSer	406	739	GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuLeu	758
QY	1174	AAATCCCTGGTCATATGTCACCTCAGAGAAATTTGATGAGACAGCAACAGGACCATG	1233	2218	GCTAGTTTATCATCTCTTATAGTAGGCTGAAACAGTTTAAACATAACCTTATGCAACA	2277
Db	407	AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValArgSerAsnArgThrValAspLeu	426	759	AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro	778
QY	1234	CTCGACTCGTAAACAATAAGCATGCAATTTATAGAAAATGCTACCTTCATTGTCCACGA	1293	2278	TTAGCAACAATAAATGCTAAACT---AATTTAGATTCAGCCCATCAACCAAGCTAAT	2334
Db	427	LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu	446	779	MetLysAlaLeuTyrProValThrThrLysAsnLeuLysLeuSerValLysAsnGlyAsn	798
QY	1294	TTGCGCGGAGATTTGCAATGCACCAATAGAAAATAAACATTAGACACACACAGCT	1353	2335	ACGGATAAACAGCATTTTGTATGATGAACACCCCAATTTTGTGAA-----	2379
Db	447	AsnLysArgSerLysArgAsnIleGluGlyGlyIleGluAsnValAsnAsnSerThrIle	466	799	SerGlyAsnAsnSerAspGlyGluGluAsnAspAsnIleAspGluGluLysLeuGln	818
QY	1354	ATTAATCGACATCGCTCTTCAATTCGCCCATCTCTCCCAATTTCTTATCATATACAA	1413	2380	-----GCATACAAGCAGCTAAAAACCACTTTAGAACACCT-----	2415
Db	467	IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln	486	819	GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnHis	838
QY	1414	ACCATATTAATCATATGTTTATGATGATTCACACGCTTGGTGGCAATTCAGCATAGA	1473	2416	-----GCTACTAACCTTGAAGCTTTGCTCA	2439
Db	487	SerHisValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysAsnLeuGlnAsnLys	506	839	LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer	858
				2440	-----TCAACTGCTTATAATCAATTCGCAAT	2466

```

Db      859 LeuLysHisArgGlyProLysTyrLysArgLeuLysasn 871
      ||| .....|||
RESULT 5
US-08-413-118-3
: Sequence 3, Application US/08413118
: Patent No. 568920
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: ATTORNEY/AGENT INFORMATION: LIMBACH, KEITH J.
: TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
: NUMBER OF SEQUENCES: 128
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
: STREET: 530 FIFTH AVENUE, 25TH FLOOR
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/413,118
: FILING DATE: 29-MAR-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/220,151
: FILING DATE: 30-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: FROMMER, WILLIAM S.
: REGISTRATION NUMBER: 25,506
: REFERENCE/DOCKET NUMBER: 454310-2670
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 840-3333
: TELEFAX: (212) 840-0712
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 879 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: N-terminal
: US-08-413-118-3

Alignment Scores:
Pred. No.: 3,33e-157 Length: 879
Score: 1968.00 Matches: 392
Percent Similarity: 62.66% Conservative: 155
Best Local Similarity: 44.90% Mismatches: 258
Query Match: 34.74% Indels: 68
DB: 1 Gaps: 10

US-09-147-052-3 (1-3261) x US-08-413-118-3 (1-879)
QY      25 ATATTTTCCTAGTATCTA----- 48
Db      8 IlePhePheIleIleIleIleIleIleCysAspProThrProGluSerThrIle 27
QY      49 -----TATGGTACGAACATCATCTCCGAGTACCCAAATGTGACATCAAGAGAA 96
Db      28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspIleArgGlu 47
QY      97 GTTGTTTCAGGCTCCAGTGTG---TCTGAGGAGAGCTACGTTTATCTTTCTCCGCCA 153
Db      48 IleLeuArgGluSerGlnIleGluSerAspThrSerThrPheTyrMetCysProPro 67
QY      154 CCAGTGGGTTCACCGGTGATCCGCTCTAGAACCCGCGGAAATGTCCCGAACCTAGAAA 213
      ||| |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu 87
QY      214 GCCACCGAGTGGGTGAAGGAATCGCGATATATTATTAAGAGATATATCAGTCCATATAAA 273
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      88 GlyLysAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleSerProTyrLys 107
QY      274 TTTAAAGTGAAGCTTTATTTATAAATATCATTCAGACGACGATCGACGGGACGACA 333
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      108 PheLysAlaAsnIleTyrLysAsnIleIleIleThrValTyrPheSerGlySerThr 127
QY      334 TATAGACAGATCACTAATCGATATACATAGAGAGCGCGCTTCATTCATTCAGACATCAG 393
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      128 TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr 147
QY      394 GATCTAATCAGCGCAAGGAAGATGCTCATCTAAAGCAAGATACCTTAAAGCAAGTGA 453
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      148 GluLeuIleAspArgArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAspTyr 167
QY      454 TATGTTGAAGCGTTTGACAGGGATGCGGGAGAAACAAAGTACTTCTTAAACCATCAAAA 513
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187
QY      514 TTCAACACGCGCGAATCTAGGGATGCGCACGACTAATGAGACGTATACCGTGGGGA 573
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      188 PheAsnThrProGlySerArgGlyTyrPheIleThrValAsnAspThrTyrThrLysIleGly 207
QY      574 TCACCATGGATATATCAACGCGAAGCTCGTCAATTGTATAGTAGAGGAATGATGCC 633
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      208 GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla 227
QY      634 CGTCTGTGTTCCGTAATCATTTTGGCAATGGCCACATCGCGACATCGGACATATCT 693
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      228 ArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspIleIleHisMetSer 247
QY      694 CCATTTTATGCTCTATCCACAGGCTGCGCGACAGCCATCGGATATATCCACGAT 753
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      248 ProPhePheGlyLeuArg---AspGlyAlaHisThrGluTyrIleSerTyrSerThrAsp 266
QY      754 AATTTCAACAACTAGATAGTATTTTCAATGATTGGCAAGCTCGAAAGCAAGC 813
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      267 ArgPheGlnGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGlnLeuGly 286
QY      814 CTTCCAGTCAAGCGTAACTTCTCATCATCATCATCATCATCATCATCATCATCATCAT 873
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      287 AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaTyrAsnTyrVal 306
QY      874 CCAAAACTACTCGTGTATCTCAATGACTAGTGAAGAGAGGTGACTGAATCTTCGT 933
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      307 ProLysIleArgGluValCysThrLeuAlaLysTyrPargGluIleAspGluIleIleArg 326
QY      934 GCAACAGTTAATGGGAGATACAGATTTATGGCGCGTGAACCTTCGGCAAGCTTTATCAGT 993
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      327 AspGluTyrLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer 346
QY      994 AATACGACTGAGTTGATCCAAATCGCATATATAGGACAATGTATTAACCGGAGCA 1053
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      347 AspThrThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla 366
QY      1054 GAAGCAGCATCAGCAGATATTTAGGACAAATAATATACATCAGTCAAGTTCAGT 1113
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      367 IleGluAlaIleAspLysIleTyrLysLysTyrAsnLysThrHisIleGlnThrGly 386
QY      1114 CATGTACAATATTTCTGGCTCTCGGGGATTTATTTAGCATATCAGCTGTCTATCC 1173
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      387 GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer 406
QY      1174 AATTCCTGGCTCATATGTACCTCAGAGATTTGATGAGACAAACAGGACCGATGAGATG 1233
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      407 AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValArgSerAsnArgThrValAspLeu 426
QY      1234 CTCGACTGGTAAACAAAGCATGCAATTTTAAAGAAAATGCTACCTCATTTGTCACGA 1293
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      427 LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu 446

```

QY 1294 TTGGGGGAGATATCGAATGACCAAAATAGAAAATAACATATAGACGACACACAGCT 1353
 Db 447 AsnLysArgSerLysArgAsnIleGluGlyGlyLeuAsnValAsnSerThrIle 466
 QY 1354 ATTAATATCGACATCGTCTGTTCAATTCGCCCATGCTCCAAATTCCTTATGATCATATACAA 1413
 Db 467 IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln 486
 QY 1414 ACCATATATATGATATGTTAGTAGATGTCACAGCTGGTGGCGGAATTCGAGATAGAGA 1473
 Db 487 SerHisValAsnGluMetLeuSerArgIleAlaThrAlaTyrCysAsnLeuGlnAsnLys 506
 QY 1474 GAATCTGTTTATGGCAGGAGGATGATTAATCTAGCTAGCGCTACAGCGAGTGCACAA 1533
 Db 507 GluArgThrLeuTrpAsnGluValMetLysLeuAsnProThrSerValAlaLaSerValAla 526
 QY 1534 TTAGGAAGGAGAGTGGTGCACAAAGATGTTGGGGATGCTGCTGTATCGAGTGCACCT 1593
 Db 527 MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal 546
 QY 1594 GCTATAGATCGGAATCCGTCACCTTTGCAAAATTCATGCGAGTATCATCATCCACTAAT 1653
 Db 547 AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr 566
 QY 1654 ACATGTTATAGCCGACCATGTTGTTCTATTTTCATATGGAGAAAACCAAGGAAC--ATA 1710
 Db 567 ThrCysTyrSerArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrIle 586
 QY 1711 CAGGGACAACCTCGTGTGAAAACACAGAGTTCCTCCACGCTAGAGGCTGTAGAGCCATGC 1770
 Db 587 GluGlyGlnLeuGlyGluAsnGlnLeuValGluArgLysLeuIleGluProCys 606
 QY 1771 TCGGCTAATCATCGTATGATATTTCTGTTGGATCCGGTTATGCTTTATTTGAAAACAT 1830
 Db 607 ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr 626
 QY 1831 AATTTGTTAGATGTTAGAGCTGCGGATATACAGATTCCTAGCAGTATGTCGAGCTT 1890
 Db 627 AlaTyrValArgLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu 646
 QY 1891 AATCTAACCTCGTGTAGAGATCGGAAATTTGCTTTATCGGTTTACACAAAAGAGAG 1950
 Db 647 AsnIleThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu 666
 QY 1951 TTGCGTGTATGTTGTTGATTTATGCAAGATAGCTCCGCCCAATCAACATACATGAA 2010
 Db 667 LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAla 686
 QY 2011 CTTAATATTTATGACATAACAAGTAATAGAGTGGATCAATATACGGCGGGCTGCAG 2070
 Db 687 LeuLysPheTyrAspIleAspSerValValLysValAspAsnValValIleMetArg 706
 QY 2071 GAATTCGGCTGTATGCTATTACTAAAGATGCAAAACCAATATATGCAAAACCCAA 2130
 Db 707 GlyIleAlaAsnPheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal 726
 QY 2131 TTAGAAGCAGCGGAATGGAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATAGCT 2190
 Db 727 LeuGlyAlaAla-----AsnAlaValIleAlaThrValSer 738
 QY 2191 TCACATCAACACTATGCCAAG-----ATGAA 2217
 Db 739 GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuLeu 758
 QY 2218 GCTAGTTTATCATCTGCTTATGAGCTGAAACAGTGAACATTAACATTAACCTTAATGCAACA 2277
 Db 759 AlaGlyPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro 778
 QY 2278 TTAGAACAACATAAATGGCTAAACCT---AATTTAGAATCAGCATCAACCAAGCTAAT 2334
 Db 779 MetLysAlaLeuTyrProValThrThrLysAsnLeuLysGluSerValLysAsnGlyAsn 798

QY 2335 ACGATAAAACGACTTTTGTATATGAACACCAACCAATTTAGTTGAA----- 2379
 Db 799 SerGlyAsnAsnSerAspGlyGluGluAsnAspAsnIleAspGluGluLysLeuGln 818
 QY 2380 -----GCATACAAACGACTAAAAACCACTTTAGAACACGCT----- 2415
 Db 819 GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnGluHis 838
 QY 2416 -----GCTACTAACCTTTGAAGTTTGTCA 2439
 Db 839 LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer 858
 QY 2440 -----TCAACTGCTTATTAATCAAAATTCGCAAT 2466
 Db 859 LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn 871
 RESULT 6
 US-08-413-118-106
 ; Sequence 106 Application US/08413118
 ; Patent No. 5688920
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/413,118
 ; FILING DATE: 29-MAR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/220,151
 ; FILING DATE: 30-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FROMMER, WILLIAM S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2670
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 106:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 879 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; US-08-413-118-106
 Alignment Scores:
 Pred. No.: 3.33e-157 Length: 879
 Score: 1968.00 Matches: 392
 Percent Similarity: 62.66% Conservative: 155
 Best Local Similarity: 44.90% Mismatches: 258
 Query Match: 34.74% Indels: 68
 DB: 1 Gaps: 10
 US-09-147-052-3 (1-3261) x US-08-413-118-106 (1-879)
 QY 25 ATATTTTCTCTATAGTATTCTA----- 48

[illegible]

367	Db	IleGluAlaIleAspLysIleTyrLYLsYsLysTyrAsnLysThrHisIleGlnThrGly	386
1114	QY	CATGTACAATATTCTTGGCTCTCGGGGGAATTTATGTAGCATATCAGCGTGTCTATCC	1173
387	Db	GlueuGluThrTyrLeuAlaAargGlyGlyPheIleIleAlaPheargProMetIleSer	406
1174	QY	AAATCCCTGGCTCATATGTACCTCAGAGAATGTATGAGAGACACAGGCCGATGAGATG	1233
407	Db	AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValAargSerAsnAargThrValAspLeu	426
1234	QY	CTGCACCTGGTAAACAATAAGCATGCAATTTATAGAANAATGCTACCTCATTTGTCACGA	1293
427	Db	LysSerLeuLeuAsnProSerValAargGlyGlyAlaAargLysAargSerValGluGlu	446
1294	QY	TTGGCGGCAGATATTCGAAATGCACCAATAGAAAAATAACATTAGACGACACACAGCT	1353
447	Db	AsnLysAargSerLysAargAsnIleGluGlyClyIleGluAsnValAsnAsnSerThrIle	466
1354	QY	ATTTAAATCGACATGCTGTGTCAATTCGCCATGCTCCAAATTTCTTTATGATCATACAA	1413
467	Db	IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln	486
1414	QY	ACCATATTATGATATGTTTAGTAGATTGCCACACGTTGGTCGGAATTCGACAGATAGA	1473
487	Db	SerHisValAsnGluMetLeuSerAargIleAlaThrAlaTyrCysAsnLeuGlnAsnLys	506
1474	QY	GAACCTTTTATGCGACGAGGATAAAGATTAACTAGCCGCTACAGCGAGTCGAACA	1533
507	Db	GluAargThrLeuTyrAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla	526
1534	QY	TTAGGAAGGAGATGGCTGCAAGATGTTGGGGATGTCGCTGCTGTATGAGCTGCAC	1593
527	Db	MetAspGlnAargValSerAlaAargMetLeuGlyAspValLeuAlaValThrGlnCysVal	546
1594	QY	GCATATAGATCGGNAATCCGTCACCTTTGCCAAAATTCATGCGAGTTATCATCATCA	1653
547	Db	AsnIleSerGlySerSerValPheIleGlnAsnSerMetAargValLeuGlySerThr	566
1654	QY	ACATGTTATAGCCGACCATTTGGTTCATTTTCATATGGAGAAAACCAAGGAAC---	1710
567	Db	ThrCysTyrSerAargProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrIle	586
1711	QY	CAGGGACAATCGGTGAAAACACAGAGTTCCTTCCACGCTAGAGGCTGTAGAGCCATGC	1770
587	Db	GluGlyGlnLeuGlyGluAsnAsnGluLeuValGluAargLysLeuIleGluProCys	606
1771	QY	TCGGCTAATCATCGTATAGATATTTCTGTTTGGATCCGGTTATGCTTTATTGGAAC	1830
607	Db	ThrAlaAsnHisLysAargTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr	626
1831	QY	AAATTTTGTAAAGTGTAGACGCTGCCGATATACAGATTCTAGACATTTGTCGAGCTT	1890
627	Db	AlaTyrValAargLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu	646
1891	QY	AATCTAACCTGCTAGNAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAAGAAG	1950
647	Db	AsnIleThrLeuLeuGluAspAargGluPheLeuProLeuGluValTyrThrAargAlaGlu	666
1951	QY	TTCCGTGATCTTGGTCTATTGGATTATGCAGAGTAGCTCGCCGCCAATCACTACATGAA	2010
667	Db	LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnAargAargAsnGlnLeuHisAla	686
2011	QY	CTTAAATTTTATGACATAAACAAGTAATAGAAGTGCATACAAATTAACGGGGCTGCAC	2070
687	Db	LeuLysPheTyrAspIleAspSerValValLysValAspAsnAsnValIleMetAarg	706
2071	QY	GAATTCCGGCTGTATGCTATTACTATAAAAAAGTCGCAACCCAAATTAATGCCCAAC	2130
707	Db	GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValCylAlaGlyPheGlyLysValVal	726
2131	QY	TTAGAAGCCGCGAATGGAGTTAAACAGATCTTAATCAATGCTTAACGGATGACATTAG	2190
727	Db	LeuGlyAlaAla-----AsnAlaValIleAlaThrValSer	738


```
Db      307  ProLysIleArgGluValCysThrLeuAlaLysTrpArgGluIleAspLulleIleArg 326
QY      934  GCAACAGTTAATGGGAGATACAGATTATGGCCCGTGAACCTTCGGCAACGGTTTATCAT 993
Db      327  AspGluTyLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer 346
QY      994  AATACGACGTAGCTTGGATCCCAATCGCATCATATTAGGACATGATTAAACGCGAGGCA 1053
Db      347  AspThrThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla 366
QY      1054  GAAGCAGCAATCGAGCAGATATTTAGGACAAAATATAATGACAGTCACGCTCAAGGTGGA 1113
Db      367  IleGluAlaIleAspLysIleTyrLysLysTyrAsnLysThrHisIleGlnThrGly 386
QY      1114  CATGTACAAATATCTTGCTCTCGGGGATTTATGTAGCATATCAGCGCTGTTCTATCC 1173
Db      387  GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer 406
QY      1174  AATCCCTGGCTCATGTACCTCAGAGNATGTAGAGACACACAGGACCGATGAGATG 1233
Db      407  AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValArgSerAsnArgThrValAspLeu 426
QY      1234  CTCGACCTGGTAAACAATAAGCATGCAATTTATAAGAAAAATGCTACCTCATTTGCACGA 1293
Db      427  LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu 446
QY      1294  TTGCGCGGAGATATTCGAATATCCACCAATAGAAAAATAACATTAGACGACACACAGCT 1353
Db      447  AsnLysArgSerLysArgAsnIleGluGlyIleGluAsnValAsnAsnSerThrIle 466
QY      1354  ATTAATATGCATCTGCTGTTCAATTCGCATGCTCCCAATTTCTTTATGATCATATACAA 1413
Db      467  IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln 486
QY      1414  ACCATATTAATGATATGTTTAGTATGTCACACGCTTGTCGGAATTCGAGAAATAGA 1473
Db      487  SerHisValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysAsnLeuGlnAsnLys 506
QY      1474  GAACCTTGTTTATGCGACCAAGGATAAAGATTATCTAGCGCTACAGCGATCGACACA 1533
Db      507  GluArgThrLeuTrpAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla 526
QY      1534  TTAGGAAGAGAGTGGCTGCAAGATGTTGGGGATGCGCTGCTGTTATCGAGCTGCAC 1593
Db      527  MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal 546
QY      1594  GCTATAGATCGGAATCCGTCATCTTGCAAAATCTATCGGAGTTATCACATCCACTAAT 1653
Db      547  AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr 566
QY      1654  ACATCTTATACCGACCATTTGTTCTATTTTCATATGAGAGAAACCAAGGAAC---ATA 1710
Db      567  ThrCysTyrSerArgProLeuIleSerPheLysAlaLeuLeuAsnSerThrAsnTyrIle 586
QY      1711  CAGGACCAACTCGGTGAAACAACAGTTGCTTCCACGCTAGAGCGCTGTAGAGCCATGC 1770
Db      587  GluGlyGlnLeuGlyGluAsnAsnGluLeuLeuValGluArgLysLeuIleGluProCys 606
QY      1771  TCGGCTAATCATCGPAGATATTTCTGTTGGATCCGGTTATGCTTTTATTTGAAACTAT 1830
Db      607  ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr 626
QY      1831  AATTTTGTAGATGTTAGATGCTGCCATATACAGATGTGTAGACATTTGTTCGAGCTT 1890
Db      627  AlaTyrValArgLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu 646
QY      1891  AATCTAACCTCGTAGAAGATCGGAAATTTGCTTTATCGGTTTATCGGTTTACACAAAAGAG 1950
Db      647  AsnIleThrLeuLeuGluAspArgPheLeuProLeuProLeuGluValTyrThrArgAlaGlu 666
QY      1951  TTGCGTGATGTTGGTATGATTCGATATCGAAGTAGCTCCCGCAATCAACTACATGAA 2010
```

```
667  LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAla 686
2011  CTTAAATTTTATCACATAAACAAGTAATAGAGTGGATACATAAATTAACGGCGCTGCAG 2070
687  LeuLysPheTyrAspIleAspSerValValLysValAspAsnValValIleMetArg 706
2071  GAATTCGGCTGTATGCTCTATTACTTAAAGAAGTGCACCAATTAATGCGCAACCCAA 2130
707  GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal 726
2131  TTAGAAGCAGCGCAATGGATTAACAGATCTTAATCAATGCTTAACCGATGATCATAGCT 2190
727  LeuGlyAlaAla-----AsnAlaValIleAlaThrValSer 738
2191  TCATACAGACTATGCCAAG-----ATTGAA 2217
739  GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeu 758
2218  GCTAGTTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACAATTAACCTTAATGCAACA 2277
759  AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro 778
2278  TTAGAACAACATAAAATGCTAAACT---AATTAGATCAGCCATCAACCAAGCTAAT 2334
779  MetLysAlaLeuTyrProValThrThrLysAsnLeuLysGluSerValLysAsnGlyAsn 798
2335  ACGGATAAAACGACTTTTGATTAATGAACACCAAAATTTAGTTGAA----- 2379
799  SerGlyAsnAsnSerAspGlyGluGluAsnAspAspAsnIleAspGluLysLeuGln 818
2380  -----GCATACAAAGCCTAAACACCTTTAGAACACGT----- 2415
819  GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGluHis 838
2416  -----GCTACTAACCTTGAAGGTTTGCA 2439
839  LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer 858
2440  -----TCACTGCTTATTAATCAAAATTCGAAT 2466
859  LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn 871

RESULT 8
US-08-473-446-3
; Sequence 3, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESS: CURTIS, MORRIS & SAFFORD, P.C.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
```

REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454310-2670
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 879 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: N-terminal
 US-08-473-446-3

Alignment Scores:

Pred. No.: 3,33e-157 Length: 879
 Score: 1968.00 Matches: 392
 Percent Similarity: 62.66% Conservative: 155
 Best Local Similarity: 44.90% Mismatches: 258
 Query Match: 34.74% Indels: 68
 DB: 3 Gaps: 10

US-09-147-052-3 (1-3261) x US-08-473-446-3 (1-879)

QY 25 ATATTTTTCCTATAGTATTCTA----- 48
 DB 8 llePhePheillellyrThrLeullelleCysAspProThrThrProGluSerThrIle 27
 QY 49 -----TATGTAGCAACTCATCTCGAGTACCCAAATGTGACATCAAGAA 96
 DB 28 AsnProLeuAsnHisHisAsnLeuSerThrProLysProThrSerAspIleArgGlu 47
 QY 97 GTTGTTCGAGCGTCCAGTTG---TCTGAGGAAGAGTCTACGTTTTTATCTTGTCCCA 153
 DB 48 lleLeuArgGluSerGlnleGluSerAspThrSerThrPheTyMetCysProPro 67
 QY 154 CCAGTGGGTTCACCGTATCGCTGTAGACCGCGCGAATAATGTCGCACTAGAAA 213
 DB 68 ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrlsLeu 87
 QY 214 GCCACCGAGTGGGTGAAGGAATCGGATATATTAAAGAGAATATCAGTCCATATAA 273
 DB 88 GlyLysAsnPheThrGluGlylleAlaValillePheLysGluAsnIleSerProTyrls 107
 QY 274 TTTAAAGTCAGCTTTTATATAAAATATCATTCAGACGACATGGACGGGGACGACA 333
 DB 108 PheLysAlaAsnIleTyrlsAsnIlelleThrValTrpSerGlySerThr 127
 QY 334 TATACAGATCATAATCGATATACATAGGACGCGCGTTTCCATTGAAGATCAGC 393
 DB 128 TyrAlaValilleThrAsnArgTyrlsAspArgValProilleGlyValProGluIleThr 147
 QY 394 GATCTAATCGCGGCAAGAGATGCTCATCAAGCAAGATACCTTAGAACAATGTA 453
 DB 148 GluLeulleAspArgGlyMetCysLeuSerLysAlaAspTyrlleArgAsnAsnTyrls 167
 QY 454 TATGTTGAAGCGTTTGACAGGATCGGCGAGAAAACAAGTACTTCTAAAACATCAAAA 513
 DB 168 GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys 187
 QY 514 TTCACACGCGCGAATCTAGGCGATGGCAGCAGTATAGACGTATACCGTGGGGGA 573
 DB 188 PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrlsIleGly 207
 QY 574 TCACCATGGATATACGACGGGAACCTCCGTCATTTCTATAGTAGAGGAATGATGCC 633
 DB 208 GlySerGlyPheTyrlsSerGlyThrSerValAsnCysIleValGluValAspAla 227
 QY 634 CGCTGTGTTTCCGTATTCATATTTTCAATGGCCAATGGCAGATGGCGAACAATATCT 693
 DB 228 ArgSerValTyrlsProTyrlsAspSerPheAlaIleSerThrGlyAspIleHisMetSer 247

QY 694 CCATTTTATGGTCTATCCACCACAGAGGCTGCCGACAGACCCATGGGATATCCCGAGAT 753
 DB 248 ProPhePheGlyLeuArg---AspGlyAlaHisThrGluTyrlsSerThrAsp 266
 QY 754 AATTTTCAAACTAGATAGCTATTTTCAATGGATTTGGCAAGGCTGCAAAAGCAAGC 813
 DB 267 ArgPheGlnIleGluGlyTyrlsProIleAspLeuAspThrArgLeuGlnIleGly 286
 QY 814 CTTCCAGTCAACGCTTCTCATCATCAGTTCACACTTCCAGTGGTGGGAGCGGCT 873
 DB 287 AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaTrpAsnTrpVal 306
 QY 874 CCAAAAACACTCTCGTGTATGTTCAATGACTAGTGAAGAGGTGACTGAAATGTTCGT 933
 DB 307 ProLysIleArgGluValCysThrLeuAlaLysTrpArgGluIleLeuArg 326
 QY 934 GCAACAGTAAATGGGAGATACAGATTTATGGCCGCGTAACTTTGGCAACGTTTATCAGT 993
 DB 327 AspGluTyrlsGlySerTyrlsArgPheThrAlaLysSerIleSerAlaThrPheIleSer 346
 QY 994 AATACGACTGAGTTTATCCAAATCGCATATATTAGGACAATGATTAAAGCGGAGCA 1053
 DB 347 AspThrThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla 366
 QY 1054 GAAGCAGCAATCGACAGATATTAGGACAAATAATATATAGCAGTACACGCTCAAGTTGA 1113
 DB 367 lleGluAlaIleAspLysIleTyrlsLysLysTyrlsAsnLysThrHisIleGlnThrGly 386
 QY 1114 CATGTACAAATTTCTGGCTCTCGGGGATTTATGTAGCATATACAGCTCTTCTATCC 1173
 DB 387 GluLeuGluThrTyrlsLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer 406
 QY 1174 AAATCCCTGGCTCATATGTACAGAAATTTAGTACAGACAGACAGACCATGAGATG 1233
 DB 407 AsnGluLeuAlaLysLeuTyrlsAsnGluLeuValArgSerAsnArgThrValAspLeu 426
 QY 1234 CTCGACCTGGTAAACAATAGCATGCAATTTATAGAAAATGCTACCTCATGTCACGA 1293
 DB 427 LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu 446
 QY 1294 TTGCGGCGAGATATCGAAATCGCAAAATACCAATAAATAATAGACACACACAGCT 1353
 DB 447 AsnLysArgSerLysArgAsnIleGluGlyGlyIleGluAsnValAsnAsnSerThrIle 466
 QY 1354 ATTAATCGACATCGTCTGTTCAATTCGCCATGCTCCATTTCTTTATGATCATATACAA 1413
 DB 467 IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrlsAspHisIleGln 486
 QY 1414 ACCCATATATGATATGTTTAGTAGGATGGCAGCTGGTGGCAATGTCAGATAGA 1473
 DB 487 SerHisValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysAsnLeuGlnAsnLys 506
 QY 1474 GAACCTGTTTATGCGACAGGAGTAAAGATTATCTAGCGCTACACGAGTGCACACA 1533
 DB 507 GluArgThrLeuTrpAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla 526
 QY 1534 TTAGGACGAGAGTGGCTGCAAGATGTTGGGGGATGTCGCTGCTGATTCGAGCTGCACCT 1593
 DB 527 MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal 546
 QY 1594 GCTATAGATGGGAATCCGTACCTTTGCAAAATCTATCGAGTATATACATCCACTAAT 1653
 DB 547 AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr 566
 QY 1654 ACATCTTATAGCGACCATTTGCTTCTATTTTCATATGGAAGAAACCAAGGAAC---ATA 1710
 DB 567 ThrCysTyrlsArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrls 586
 QY 1711 CAGGACAACTCGGTGAAACAACAGTTCGTTCCACCGCTAGAGCGCTAGAGCATGCTC 1770
 DB 587 GluGlyGlnLeuGlyGluAsnAsnGluLeuValGluArgLysLeuIleGluProCys 606
 QY 1771 TCGGCTAATCATCGTAGATATTTTCTGTTGGATCCGGTTATGCTTTATTTGAAACATAT 1830

Qy	514	TTCAACACGCCCGAATCTAGGCGATGCCACACGACTAATGAGAGCGTATACCGTGTGGGA	573
Db	188		
Qy	574	TCACCATGGATATCATCCAGCGGAACCTCGGTCAATTGTATAGTAGAGGAAATGGATGCC	633
Db	208		
Qy	634	CGCTCTGTGTTTCGGTATTCTATATTTTGCATATGGCCCAATGCGCAGACATCGGAACATATCT	693
Db	228		
Qy	694	CCATTTTATGGTCTATCCACACAGAGCGTCCGCCACAACCCATGGGATATCCCCAGGAT	753
Db	248		
Qy	754	AATTTCAAAACAACCTAGATAGTATTTTTCATGGATTGGACAACGCTCGAAAGCAAGC.	813
Db	267		
Qy	814	CTTCCAGTCAAGCGTAACCTTCTCATCATACATCACACTTCACAGTTGGGTGGGACTGGGCT	873
Db	287	AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaTspAsnTspVal	306
Qy	874	CCAAAACTACTCGTGTATGTTCAATCACATAAGTGGAAAGAGGTGACTGAAATGTTCCGT	933
Db	307	ProLysIleArgGluValCysThrLeuAlaLysTspArgGluIleAspGluIleIleArg	326
Qy	934	GCAACAGTTAATGGGAGATACAGATTTATGGCCGCTGAACCTTCGCGCAACGTTTATCAGT	993
Db	327	AspGluThrLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer	346
Qy	994	AATAGACTGAGTTTGATCCAAATCGCATCATATTAGACAAATGTATTAACCGGAGGCA	1053
Db	347	AspThrThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla	366
Qy	1054	GAAGCACAATCAGACAGATATTAGACAAAAATATATCACATGCACGTCACGTCGAGTTGGA	1113
Db	367	IleGluAlaIleAspLysIleTyrLysLysTyrAsnLysThrHisIleGlnThrGly	386
Qy	1114	CATGTACAATATTTCTTGGCTCTCGGGGATATTATTAGCATCATACGCTGTCTATCC	1173
Db	387	GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer	406
Qy	1174	AAATCCCTGGCTCATATGTACCTCAGAGATTTGATGAGACACAACAGGACCGATGAGATG	1233
Db	407	AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValArgSerAsnArgThrValAspLeu	426
Qy	1234	CTCGACCTGCTAAACAATAAGCATGCAATTTATAAGAAAAATGCTACCTCATTTGTCACGA	1293
Db	427	LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu	446
Qy	1294	TTGGCGCGAGATATTTCGAAATGCAACCAATAGAAAAATAACATTAGACGACACACCACT	1353
Db	447	AsnLysArgSerLysArgAsnIleGluGlyGlyIleGluAsnValAsnAsnSerThrIle	466
Qy	1354	ATTAATATCGACATCGTCTGTTCATTCGCCCATGCTCCAATTTCTTTATGATCATATACAA	1413
Db	467	IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln	486
Qy	1414	ACCATATTATGATATGTTTGTAGGATGGCCACAGCTTGGTCGGAATTCAGACATAGA	1473
Db	487	SerHisValAsnGluMetLeuSerArgGlyAlaThrAlaTspCysAsnLeuGlnAsnLys	506
Qy	1474	GAACCTGTTTATGGCAAGGAGTAAGATTAATCTACGCGCTACAGCGAGTCAACA	1533
Db	507	GluArgThrLeuThrAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla	526
Qy	1534	TTAGGAAGGAGATGGCTGCAAGATGTTGGGGATGTCGCTGCTATCGAGCTGCAC	1593
Db	527	MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal	546

RESULT 10

RESULT 10
US-08-525-742-10

US-08-323-742-10
: Sequence 10, Application US/08525742

; Patent No. 5871742

GENERAL INFORMATION:

QY	1594	GCATATAGATCGCGAATCCCGTCACTCTTGGTGCAGAAATTCATGCGAGTTATPCACATCCCACTAAT	1655
Db	547	AsnIleSerGlySerValPheIleGlnAsnSerMetArgValLeuGlySerThr	566
QY	1654	ACATGTTATAGCGACCATGGTTCATATTTTCATATGGAGAAACCAAGAAAC---ATA	1710
Db	567	ThrCysTyrSerArgProLeuIleIleSerPheLysAlaLeuGlnAsnSerThrAsnTyrIle	586
QY	1711	CAGGGACAACCTCGGTGAAAAACAACGAGTTTCCTTCCAAACGCTAGAGCGCTGTAGAGCCATGC	1770
Db	587	GlucylGlnLeuGlyGluAsnAsnGluLeuLeuValGluArgLysLeuIleGluProCys	606
QY	1771	TCGCGTAATCATCGTAGATATTTTCGTTTGGATCCGTTATCGTTATTTGTAAGAACTAT	1830
Db	607	ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr	626
QY	1831	AAATTTGTTAAGATGTTAGACGCTGCCGATATACAGATTCCTACGACATTTGTCGAGCTT	1890
Db	627	AlaTyrValArgLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu	646
QY	1891	AACTAACCTGCTAGAAAGATCGGGAAATTTTGCCTTTATCCGTTTACACAAAGAAGAC	1950
Db	647	AsnIleThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu	666
QY	1951	TTGCGTCATCTGGTGTTATGGATTATGCAAGATAGCTCGCGCCATCACTCACTACATGAA	2010
Db	667	LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAla	686
QY	2011	CTTAATATTTTACACATAAACAAGTAATAGAAGTGTGATACAAATTCACGGCGCTGCAG	2070
Db	687	LeuLysPheTyrAspIleAspSerValValLysValAspAsnAsnValIleMetArg	706
QY	2071	GAATTCGGCTGTATGCTATTACTATAAAGATGCAACCCCAATATATGCCAAACCCAA	2130
Db	707	GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal	726
QY	2131	TTAGAAGCAGCGGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCT	2190
Db	727	LeuGlyAlaAla-----AsnAlaValIleAlaThrValSer	738
QY	2191	TCACTACAGACTATGCCAAG-----ATTGAA	2217
Db	739	GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeu	758
QY	2218	GCATGTTTATCATCTGCTTTATAGTGAAGCTTAACAGTTAACCAATACCTTATGCAACA	2277
Db	759	AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLysLeuLysSerAsnPro	778
QY	2278	TTAGAACAACATAAAATGGCTAAAACT---AATTTAGAATPCAGCCATCAACCAAGCTAAT	2334
Db	779	MetLysAlaLeuTyrProValThrThrLysAsnLeuLysGluSerValLysAsnGlyAsn	798
QY	2335	ACGGATAAACGACTTTTGATATGAAACACCCCAATTTAGTTGAA-----	2379
Db	799	SerGlyAsnAsnSerAspGlyGluGluAsnAspAspAsnIleAspGluGluLysLeuGln	818
QY	2380	-----GCATACAAAGCACTAAACACCTTTAGAACACCT-----	2415
Db	819	GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnGluHis	838
QY	2416	-----GCTACTACCTTCAAGGTTTGTGCA	2439
Db	839	LysAlaIleLysLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer	858
QY	2440	-----TCACTGCTTATATCAATTCGCAAT	2466
Db	859	LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn	871

Sequence 8, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsawa, Ikuroh
APPLICANT: Funato, Hirono
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemasa
APPLICANT: Takahashi, Kiyoohito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
TITLE OF INVENTION: AS USE THEREOF
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-8

Alignment Scores:
Pred. No.: 4.23e-148 Length: 610
Score: 1859.00 Matches: 366
Percent Similarity: 96.68% Conservative: 13
Best Local Similarity: 93.37% Mismatches: 0
Query Match: 32.82% Indels: 0
DB: 2 Gaps: 0

US-09-147-052-3 (1-3261) x US-08-525-742-8 (1-610)

QY 2050 ACAATATACGGGGCTGCAGAAATCGCTGTATGCTATTAATAAAGATGCAAAAC 2109
|||||
Db 17 ThrSerPheLeuSerIleGlyLeuSerCysMetSerIleThrLysLysAspAlaAsn 36
|||||

QY 2110 CCAATATATGCGGCGGCTGAGTAAACAGATCTAATCAAT 2169
|||||

Db 37 ProAsnAsnGlyGlnThrGlnLeuGluAlaAlaArgMetGluLeuThrAspLeuLeuLeu 56
QY 2170 GCTAAAGCGATGACATTAGCTTCACTCAAGACTATGCCAAGATTGAAGCTTATCA 2229
|||||
Db 57 AlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSer 76
QY 2230 TCTGCTTATAGTGAAGCTGAACAGTTAACATTAACCTTAATGCAACATTAGAACAATA 2289
|||||
Db 77 SerAlaTyrSerGluAlaGluThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeu 96
QY 2290 AAAATGGCTAAACCTAATTTAGAAATCAGCCATCAACCAAGCTAATACGATGATAAAGCT 2349
|||||
Db 97 LysMetAlaLysThrAsnLeuGluSerAlaLysLeuGlnAlaAsnThrAspLysThrThr 116
QY 2350 TTTGATATGAACACCCCAATTTAGTTGAAGCATACAAAGCACATAAAACCCACTTTAGAA 2409
|||||
Db 117 PheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeuLysThrThrLeuGlu 136
QY 2410 CAACGTGCTACTAACCTTGAAGTTTGTTCATCACTGCTTTATATCAATTCGCAATTAAT 2469
|||||
Db 137 GlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsn 156
QY 2470 TTAGTGGATCTATCAATAAAGCTAGTAGTTTAATACTAAACACTAGATCCACTAAT 2529
|||||
Db 157 LeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThrLeuAspProLeuAsn 176
QY 2530 GGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAGAATATTAATATACG 2589
|||||
Db 177 GlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThr 196
QY 2590 TTATCAACTATTAATGAACAAAGACTAATGCTGATGCATATCTAATAAGTTTTATTAAA 2649
|||||
Db 197 LeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuAlaAsnSerPheIleLys 216
QY 2650 AAAGTGATTCAAAATAATGAACAAAGTTTGTAGGACTTTTACAAACGCTAATGTTCAA 2709
|||||
Db 217 GluValIleGlnAsnAsnLysGlnSerPheValGlyMetPheThrAsnThrAsnValGln 236
QY 2710 CCTTCAAACTACAGTTTGTGCTTTTAGTGTGATGTAACACCCGCTCAATATATAATAT 2769
|||||
Db 237 ProSerAsnTyrSerPheValAlaPheSerAlaAspValThrProValAsnTyrLysTyr 256
QY 2770 GCAAGAAGGACGGTTTGAATGGTGATCAACCTTCAAGTAGAATTCCTGCAACACGAA 2829
|||||
Db 257 AlaArgArgThrValTrpAsnGlyAspGluProSerSerArgIleLeuAlaAsnThrAsn 276
QY 2830 AGTATCACAGATGTTCTTGGATTATAGTTTAGCTGGAACAAACACGAAGTACCATAAT 2889
|||||
Db 277 SerIleThrAspValSerTrpIleTyrSerLeuSerGlyThrAsnThrLysTyrGlnPhe 296
QY 2890 AGTTTACAACTATGGTCCATCACTGGTTATTTATATTTCCCTTATAAGTTGTTAAA 2949
|||||
Db 297 SerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLysLeuValLys 316
QY 2950 GCAGCTGATGCTAATAACGTTGGATTACAAATAATAATAATGGAATTTTCAACAA 3009
|||||
Db 317 AlaAlaAspAlaSerAsnValGlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnPro 336
QY 3010 GTTGAGTTTGCACCTTCAACTAGTCGCAATATATCTACAGCTAATCCCACTCCAGCAT 3069
|||||
Db 337 ValGluPheAlaThrSerThrSerAlaAsnAsnThrThrAlaAsnProThrProAlaVal 356
QY 3070 GATGAGATTAAAGTTGCTTAAATCGTTTTATCAGGTTTAAAGTTTGGCCAAACACAATC 3129
|||||
Db 357 AspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPheGlyGlnAsnThrIle 376
QY 3130 GAATTAAGTGTTCACCGGTGAAGGAATATGATAAAGTTGGCCCAATGATTTGGCAAC 3189
|||||
Db 377 GluLeuSerValProThrGlyGluArgAsnMetAsnLysValAlaProMetIleGlyAsn 396
QY 3190 ATTTATCTTAGCTCAATGAAATTAATGCTGATAAG 3225
|||||
Db 397 MetTyrIleThrSerSerAsnAlaGluAlaAsnLys 408

RESULT 12
US-08-220-151-4
; Sequence 4, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-4
Alignment Scores:
Pred. No.: 6.55e-147 Length: 1041
Score: 1846.00 Matches: 391
Percent Similarity: 54.21% Conservative: 163
Best Local Similarity: 38.26% Mismatches: 262
Query Match: 32.59% Indels: 206
DB: 1 Gaps: 15
US-09-147-052-3 (1-3261) x US-08-220-151-4 (1-1041)
QY 7 TATTTTACGGGAATTCGATATT----- 30
DB 22 TyrPheArgGlnArgCysPheProSerLeuLeuGlyIleAlaIaThrGlySerArg 41
QY 31 -----TTCTTATAGTT 42
DB 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCCTATAT----- 51
DB 62 ValLeuPheLeuValGlyProArgProValCluGlyGlnSerGlySerThrSerGluGln 81
QY 52 -----GGTAGCACTCATCTCCGAGTACCCAA 78
DB 82 ProArgArgThrValAlaIaThrProGluValGlyThrProProLysProThrThrAsp 101
QY 79 AATGTGACATCAAGAGAAGTTGTCGAGCGTCCTGCTGAGGAAGAG----- 129

Db 102 ProThrAspMetSerAspMetArgGluAlaLeuArgAlaSerGlnIleGluAlaAsnGly 121
QY 130 ---TCTAGCTTTTATCTTTGTCCTCCACAGTGGTTCAACCGTGCATCCGCTAGAACCG 186
Db 122 ProSerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluPro 141
QY 187 CCGCAAAATGTCCTCCGAACCTAGAAACCCACCGAGTGGGTGGGTGAAGGAATCGCATATTA 246
Db 142 ProArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValile 161
QY 247 TTTAAAGAGATATACAGTCCATATAAATTTAAAGTGACGCTTTATTTATAAAATATCATTT 306
Db 162 PheLysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIlelle 181
QY 307 CAGACGACGATGACGCGGACGACATATAGACAGATCATAATCGATATACAGATAGG 366
Db 182 MetThrThrValTrpSerGlySerSerTyrAlaValThrThrAsnArgTyrThrAspArg 201
QY 367 ACGCCGCTTCCATTGAAGAGATCAGGATCTAATCGACGCGCAAGGAGAGTGCATCT 426
Db 202 ValProValLysValGlnIleThrAspLeuIleAspArgA-gglyMetCysLeuSer 221
QY 427 AAAGCAAGATACCTTAGAAACAATATATATGTGAAGCGTTTGACAGGGATCGGGAGAA 486
Db 222 LysAlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspPro 241
QY 487 AAACAAGTACTTCTAAACCACTCAAAATTCACACGCCCGCAATCTAGG----- 534
Db 242 ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTrpHisThr 261
QY 534 ----- 534
Db 262 TyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGly 281
QY 534 ----- 534
Db 282 SerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGlu 301
QY 534 ----- 534
Db 302 IleThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsn 321
QY 534 ----- 534
Db 322 AsnHisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAla 341
QY 535 -----GCATGGCACACGACTAATGAGACGATACCGTG 567
Db 342 SerLysTyrAsnSerValGlySerLysAlaTrpHisThrThrAsnGluThrTyrThrLys 361
QY 568 TGGGGATCACCATTGGATATATCGAACGGGAACCTCGTCAATGTATGTATAGTACGAGAAATG 627
Db 362 IleGlyAlaAlaGlyPheHisHisSerGlyThrSerValAsnCysIleValGluGluVal 381
QY 628 GATGCCGCTCTGCTGTTTCCTGTTTCAATGCGCAATGGCAGATCGCGAAC 687
Db 382 AspAlaArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspValIleHis 401
QY 688 ATATCTCCATTTTATGGTCTATCCACCAGAGGCTGCCGACAGACCCATGGATATCCC 747
Db 402 MetSerProPhePheGlyLeuArg---AspGlyAlaHisValGluHisThrSerTyrSer 420
QY 748 CAGGATAAATTCAAACAACATAGATAGCTATTTTCAATGGATTGGACAGCGCTCAAAA 807
Db 421 SerArgPheGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGln 440
QY 808 GCAAGCTTCCAGCTCAAGCTAACTTCTCATCATCATCATCATCATCATCATCATCATCAT 867
Db 441 LeuGlyAlaProValSerArgAsnPheLeuGluThrProHisValThrValAlaIaTrpAsn 460
QY 868 TGGGCTCCAAAACTACTCGTGTATGTTCAATGACTAAGTGAAGAGAGGTGACTGAATG 927
Db 461 TrpThrProLysCysGlyArgValCysThrLeuAlaLysTrpArgGluIleAspGluMet 480

QY 928 TTGGTCCACAGTAAATGGGAGATACAGATATTATGCGCCGCTGAACATCTTCGGCAACGCTTT 987
Db 481 LeuArgAspGluTyrGlnGlySerTyrArgPheThrValLysThrIleSerAlaThrPhe 500
QY 988 ATCAGTAATACGACTGAGTTGATCCAAATCGCATATATTAGACAATATTAATAACGC 1047
Db 501 IleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeuGlyAspCysAlaThrLys 520
QY 1048 GAGCGAGACGACCAATCGACGAGATATTAGCAAAATATATAGACAGTCAGCTCAAG 1107
Db 521 GluAlaAlaGluAlaIleAspArgIleTyrLysSerLysTyrSerLysThrHisIleGln 540
QY 1108 GTTGGACATGACATATTTCTGGCTCGGGGATTTATGTAGCATATACGCTGTT 1167
Db 541 ThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMet 560
QY 1168 CTATCAAAATCCCTGGCTCATATGCTCAGAGAATTGATGAGACACACAGACCGAT 1227
Db 561 IleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeuAlaArgSerAsnArgThrVal 580
QY 1228 GAGATGCTGACCTGTTAAACAATAAGCATGCAATTTATAGAAATAATGCTACCTCATG 1287
Db 581 AspLeuSerAlaLeuLeuAsnProSerGlyThrValGlnArgThrArgGlySerVal 600
QY 1288 -----TCAGATTGCGCGGAGATATTCGAAATGCACCAATAGAAAAATA---ACATTA 1338
Db 601 ProSerAsnGlnHisIleArgSerArgSerThrIleGluGlyGlyIleGluThrVal 620
QY 1339 GACGACACCAAGCATTAATAATGACATGCTGCTTCAATTCGCGCATCTCCCAATTCCT 1398
Db 621 AsnAsnAlaSerLeuLeuLysThrSerValGluPheAlaMetIleGlnPheAla 640
QY 1399 TATGATCATATACAAACCCATTAATGATATGTTTAGTAGGATGCGACAGCTTGTC 1458
Db 641 TyrAspTyrIleGlnAlaHisValAsnGluMetLeuSerArgIleAlaThrAlaTyrCys 660
QY 1459 GAATTGAGATAGACAACTGTTTATGGCAGCAGGATTAAGATTAATCTAGGCT 1518
Db 661 ThrLeuGlnAsnArgGluHisValLeuThrTrpThrLeuLysLeuAsnProGlyGly 680
QY 1519 ACAGCGAGTGAACATAGGAAGAGAGTGGCTCAAGATGTTGGGGGATGCTCGCT 1578
Db 581 ValValSerMetAlaLeuGluArgValSerAlaArgLeuLeuGlyAspAlaValAla 700
QY 1579 GTATCGAGCTGCATGATAGATGCGGAATCGCTCACTTTGCAAAATTTCTATCGAGTT 1638
Db 701 ValThrGlnCysValAsnIleSerSerGlyHisValTyrIleGlnAsnSerMetArgVal 720
QY 1639 ATCATCCACTAATACATGTTATAGCGGACCATGTTCTATTTTCATATGGAGAAAC 1698
Db 721 ThrGlySerSerThrThrCysTyrSerArgProLeuValSerPheArgAlaLeuAsnAsp 740
QY 1699 CAAGGAACATACAGGACCACTCGGTCAAAACACGATGCTTCCCAACGCTAGAGCT 1758
Db 741 SerGluTyrIleGluGlnLeuGlyGluAsnAsnAspLeuValGluArgLysLeu 760
QY 1759 GTAGAGCATGCTCGGTATATCATCGTAGATATTTCTGTTGGATCGGTATGCTTTA 1818
Db 761 IleGluProCysThrValAsnAsnLysArgTyrPheLysPheGlyAlaAspTyrValTyr 780
QY 1819 TTTGAAACTATAATTTGTTAAGATGTTAGACGCTGCGGAAATTTGCTTTATCGCTTAC 1878
Db 781 PheGluAspTyrAlaTyrValArgLysValProLeuSerGluIleGluLeuIleSerAla 800
QY 1879 TTTGTCGAGCTTAATCTAACCTGCTAGAGATCGGGAATTTGCTTTATCGCTTAC 1938
Db 801 TyrValAspLeuAsnLeuThrLeuLeuAspArgGluPheLeuProLeuGluValTyr 820
QY 1939 ACAAAGAAGAGTTGCGGTGATGTTGTTATGATTATGAGAGTAGCTCGCGCAAT 1998
Db 821 ThrArgAlaGluLeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsn 840

QY 1999 CAACTACATCAACTAAATTTTATGACATAAAACAAGTAATAGAGTGGATACAAAT--- 2055
Db 841 GlnLeuHisAlaLeuLysPheTyrAspIleAspSerIleValArgValAspAsnLeu 860
QY 2056 -----TACGCGGGCTGCAGGAATTCGGCTGT----- 2082
Db 861 ValIleMetArgGlyMetAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPhe 880
QY 2083 -----ATGCTATTACTAAAAAGATGCA 2106
Db 881 GlyLysValValLeuGlyAlaAlaSerAlaValIleSerThrValSerGlyValSerSer 900
QY 2107 AACCCAATATGCCCCAAACCAATTAGACACCGGCAATGGAGTTAACAGATCTAATC 2166
Db 901 PheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleVal 920
QY 2167 AATGCT----- 2181
Db 921 AlaAlaPheLeuAlaTyrArgTyrIleSerArgLeuArgAlaAsnProMetLysAlaLeu 940
QY 2182 -----ACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCT 2232
Db 941 TyrProValThrThrArgAsnLeuLysGlnThrAlaLysSerProAlaSerThrAlaGly 960
QY 2233 GCTTATAGTGAAGCTGAACAGTTAACAAATACCTTAATGCAACATTAGAACACATAAA 2292
Db 961 GlyAspSerAspProGlyValAspAspPheAspGluGluLysLeuMetGlnAlaArgGlu 980
QY 2293 ATGCTAAA---ACTAATTTAGATAGCCCATCAACCAAGCTAATACGGATAAACAGACT 2349
Db 981 MetIleLysTyrMetSerLeuValSerAlaMetGluGln---GlnGluHisLysAlaMet 999
QY 2350 TTTGATATATGAACACCCCAATTTAGTTGAACATAC-----NAAGCACTAAAAACC 2400
Db 1000 LysLysAsnLysGlyProAlaIleLeuThrSerHisLeuThrAsnMetAlaLeuArgArg 1019
QY 2401 -----ACTTTAGAACACGCTGCTACTAACCTTGAAGTTTGTATCACTGCTTATAT 2454
Db 1020 ArgGlyProLysTyrGlnArgLeuAsnAsnLeuAspSerGlyAspThrGluThrAsn 1039
QY 2455 CAAATT 2460
Db 1040 LeuVal 1041
RESULT 13
US-08-413-118-4
; Sequence 4, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994

242	ArgGluLeuProLeuLysProSerLysPheAsnThrProGlnSerArgGlyTrpHisThr	261
534	-----	534
262	TyrLysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaTrpAlaGly	281
534	-----	534
282	SerSerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGlu	301
534	-----	534
302	IleThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsn	321
534	-----	534
322	AsnHisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAla	341
535	-----GCATGGCACAGCACTAATGAGACCTATACCGTG	361
342	SerLysTyrAsnSerValGlySerLysAlaTrpHisThrAsnGluThrTyrThrLys	361
568	TGGGATCACCATGGATATATCGAAGCGGAACCTCCGCAATTTGTATAGAGAAATG	627
362	IleGlyAlaAlaGlyPheHisSerGlyThrSerValAsnCysIleValGluGluVal	381
628	GATGCCCGCTCTGTTCCGTATTCATATTTGGCAATGGCCATGGCGACATCCGCAAC	687
382	AspAlaArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspValIleHis	401
688	ATATCTCCATTTTATGGTCTATCCCCACAGAGGCTGGCGAGAACCCATCGGATATCCC	747
402	MetSerProPheGlyLeuArg--AspGlyAlaHisValGluHisThrSerTyrSer	420
748	CAGGATAAATTTCAACAACTAGATGACTATTTTCAATGGATTTGGCAACCGTCGAAA	807
421	SerAspArgPheGlnGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGln	440
808	GCAAGCCTTCCAGTCAAGCGTAACCTTCATCATCATCACATTCACAGTTTGGGTGGGAC	867
441	LeuGlyAlaProValSerArgAsnPheLeuGluThrProHisValThrValAlaTrpAsn	460
868	TGGGCTCCAAAACACTACTCGTGTATGTTCAATGACTAAGTCGGAAGAGGTGACGTGAATG	927
461	TrpThrProLysCysGlyArgValCysThrLeuAlaLysTrpArgGluIleAspGluMet	480
928	TTGCGTGCACACGTTAATGGAGATACAGATTTATGGCCGCTGAACTTTCCGCAACGTTT	987
481	LeuArgAspGluTyrGlnGlySerTyrArgPheThrValLysThrIleSerAlaThrPhe	500
988	ATCAGTAATACGACTGAGTTTGATCCAAATCCGATCATATTAGGACAATATTAAACGC	1047
501	IleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeuGlyAspCysAlaThrLys	520
1048	GAGCGAAGCACCAATCGAGAGATATTTAGCAACAAATAATAGACAGTCAGCTCAAG	1107
521	GluAlaAlaGluAlaIleAspArgIleTyrLysSerLysTyrSerLysThrHisIleGln	540
1108	GTTGGCATGTACAATATTTCTGGCTCGGGGGATTTATTGTAGCATATACGCTGTT	1167
541	ThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMet	560
11168	CTATGCCAATCCCTGGCTCATATGTACCTCAGAGAAATGTAGTGAGAGACACAGACCGAT	12271
561	IleSerAsnGluLeuAlaLysLeuTyrIleAsnGluLeuAlaArgSerAsnArgThrVal	580
1228	GAGATGCTCGACCTGGTAAACAATAAGCATGCAATTTTATAGCAAAATGCTACCTCATTT	12871
581	AspLeuSerAlaLeuLeuAsnProSerGlyGluThrValGlnArgThrArgGlySerVal	600
1288	-----TCACAGTGGCGGAGATATTCGAATGCACCAATAGAAAAATA---ACATTA	13381
601	ProSerAsnGlnHisArgSerArgSerThrIleGluGlyGlyIleGluThrVal	620

QY 1339 GAGCAGACACACGCTATTAAATCGACATCGCTCTGTTCAATTCGCCAGTCTCCCAATTCCTT 1398
 Db 621 ASNAASALSerLeuLeuLysThrThrSerValGluPheAlaMetIleGlnPheAla 640
 QY 1399 TATGATCATATACAAACCCATATATATGATATATTTAGTAGGATTCGCCAGCTTGGTGC 1458
 Db 641 TyrAspTyrIleGlnAlaHisValAsnGluMetLeuSerArgIleAlaThrAlaTrpCys 660
 QY 1459 GAATTCGAGTAATAGACACTGTTTATGACGACGAAGGGATAAAGATTAATCCTCAGCGCT 1518
 Db 661 ThrLeuGlnAsnArgGluHisValLeuTrpThrGluThrLeuLysLeuAsnProGlyGly 680
 QY 1519 ACAGCGAGTCCACATTAGGAAGAGAGTGGCTGCGCAAGATGTTGGGGATGCGCTGCT 1578
 Db 681 ValValSerMetAlaLeuGluArgValSerAlaArgLeuLeuGlyAspAlaValAla 700
 QY 1579 GTATCGAGTGCCTGCTATAGATCGGGAATCGCTCACTTTCGCAAAATTCATCGCAGTT 1638
 Db 701 ValThrGlnCysValAsnIleSerSerGlyHisValTyrIleGlnAsnSerMetArgVal 720
 QY 1639 ATCACAATCCACTAATATGTTATGACCGACCATGTTCTTATTTTCATATGAGAAAAC 1698
 Db 721 ThrGlySerSerThrThrCysTyrSerArgProLeuValSerPheArgAlaLeuAsnAsp 740
 QY 1699 CAAGGAACATACAGGACACACTCGGTGAAACACAGAGTGGTTCACAGCTAGAGGCT 1758
 Db 741 SerGluTyrIleGluGlnLeuGlyGluAsnAsnAspLeuValGluArgLysLeu 760
 QY 1759 GTAGAGCGATCGCTGCTATCATCGTAGATATTTCTGTTGTCGATCCGGTATGCTTTA 1818
 Db 761 IleGluProCysThrValAsnAsnLysArgTyrPheLysPheGlyAlaAspTyrValTyr 780
 QY 1819 TTTGAAACATATAATTTTGTAGATGGTAGACGCTGCCGATATACAGATTTGTAGCAC 1878
 Db 781 PheGluAspTyrAlaTyrValArgLysValProLeuSerGluIleGluLeuSerAla 800
 QY 1879 TTTGTCGAGTTAATCTAACCTGCTAGAGATCGGGAATTTTGGCTTTATCCGTTTAC 1938
 Db 801 TyrValAspLeuAsnLeuThrLeuLeuGluAspArgPheLeuProLeuGluValTyr 820
 QY 1939 ACAAAGAGAGTGGTGTGATGTTGTTATGATTATGACAGATGAGTCCGCCCAAT 1998
 Db 821 ThrArgAlaGluLeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsn 840
 QY 1999 CAACATCATGAATTAATTTTATGACATAAACAAGATTAAGAGTGGATCAAAAT --- 2055
 Db 841 GlnLeuHisAlaLeuLysPheTyrAspIleAspSerIleValArgValAspAsnAsnLeu 860
 QY 2056 -----TACCGCGGCTCGCAGGAATTCGCTCT----- 2082
 Db 861 ValIleMetArgGlyMetAlaAsnPheGlnGlyLeuGlyAspValGlyAlaGlyPhe 880
 QY 2083 -----ATGCTATTACTAAAGGATGCA 2106
 Db 881 GlyLysValValLeuGlyAlaAlaSerAlaValIleSerThrValSerGlyValSerSer 900
 QY 2107 AACCCAAATAGGCCAAACCCCAATAGAGCAGCGCAATGGAGTGAACAGATCTAATC 2166
 Db 901 PheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleVal 920
 QY 2167 AATGCT-----AAAGCGATG 2181
 Db 921 AlaAlaPheLeuAlaTyrArgTyrIleSerArgLeuArgAlaAsnProMetLysAlaLeu 940
 QY 2182 -----ACATTAGCTTCACTACAGACTAGTCCCAAGATTCAGCTAGTTTATCATCT 2232
 Db 941 TyrProValThrThrArgAsnLeuLysGlnThrAlaLysSerProAlaSerThrAlaGly 960
 QY 2233 GCTTATAGTGAAGCTGAAACAGTAAACAATAACCTTAATGCAACATTAGAACAACTAAA 2292
 Db 961 GlyAspSerAspProGlyValAspPheAspGluLysLeuMetGlnAlaArgGlu 980

QY 2293 ATGCTAAA---ACTAATTTAGATCAGCCATCAACCAAGCTATACGGATAAACGACT 2349
 Db 981 MetIleLysTyrMetSerLeuValSerAlaMetGluGln---GlnGluHisLysAlaMet 999
 QY 2350 TTTGATAATCAACACCAAAATTTAGTTGAAGCATAC-----AAAGCACTAAACACC 2400
 Db 1000 LysLysAsnLysGlyProAlaIleLeuThrSerHisLeuThrAsnMetAlaLeuArgArg 1019
 QY 2401 -----ACTTTAGAACCAAGCTGCTACTTAACCTTTGAAGCTTTTGTTCATCACTGCTTATAAT 2454
 Db 1020 ArgGlyProLysTyrGlnArgLeuAsnAsnLeuAspSerGlyAspAspThrGluThrAsn 1039
 QY 2455 CAAATT 2460
 Db 1040 LeuVal 1041
 RESULT 14
 US-08-473-446-4
 ; Sequence 4, Application US/08473446
 ; Patent No. 6017542
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/473,446
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/413,118
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FROMMER, WILLIAM S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2670
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1041 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: N-terminal
 ; US-08-473-446-4
 Alignment Scores:
 Pred. No.: 6,55e-147 Length: 1041
 Score: 1846.00 Matches: 391
 Percent Similarity: 54.21% Conservative: 163
 Best Local Similarity: 38.26% Mismatches: 262
 Query Match: 32.59% Indels: 206
 DB: 3 Gaps: 15
 US-09-147-052-3 (1-3261) x US-08-473-446-4 (1-1041)
 QY 7 TATTTAGCGGAATTCATATTT----- 30


```
QY 886 CGTGTATGTTCAATGACTAAAGTGAAGAGTGAAGTGAAGTGGT---GCAACAGTT 942
Db 372 ArgValCysSerLeuAlaLysIrpArgGluAlaGluMetThrArgSpGluThrArg 391
QY 943 AATGGAGATACAGATTTATGGCCGCGTGAAGTTCGGCAACGCTTTATCAGTAAATAGACT 1002
Db 392 AspGlySerPheArgPheThrArgAlaLeuGlyAlaSerPheValSerAspValThr 411
QY 1003 GAGTTTGATCCCAATCGCATCATATAGGACACATGATTAACGGCGGAGCAGACGCA 1062
Db 412 GluLeuAspLeuGlnArgValHisLeuGlyAspCysValLeuArgGluAlaSerGluAla 431
QY 1063 ATCGACACATATTTAGGACAAATATATGACAGTCAGTCAGTCAAGGTTGGA---CATGTA 1119
Db 432 IleAspAlaIleTyrArgArgIrpArgIrpArgIrpArgIrpArgIrpArgIrpArg 451
QY 1120 CAATATTTCTGCTCTCGGGGATTTATTTAGCATATATGATGATGATGATGATGATGAT 1179
Db 452 GluValTyrLeuAlaArgGlyGlyPheValValAlaPheArgProLeuIleSerAsnGlu 471
QY 1180 CTGGCTCATATGCTACTCAGAAATGATGAGAGACACAGACCGATGAGATGCTGCAG 1239
Db 472 LeuAlaGlnLeuTyrAlaArgGluLeu-----GluArgLeuGly 484
QY 1240 CTGGTAACAATAAGCATGCAATTTATAGAAAATGCTACCTCATTTGTCACGATTCGG 1299
Db 485 LeuAlaGly-----ValValGlyProAlaAlaProAlaAlaAlaAlaArgAlaArg 501
QY 1300 CGAGAT-----ATTCGAAATGCACCAATAGA 1326
Db 502 ArgSerProGlyProAlaGlyThrProGluProProAlaValAsnGlyThrGlyHisLeu 521
QY 1327 AAAATAACATTAGACGACACACAGCTATTAAATCGACATCTGCTGCTCAATTCGCCATG 1386
Db 522 ArgIleThr-----ThrGlySerAlaGluPheAlaArg 532
QY 1387 CTCCAATTTCTTATGATCATATACAAACCCATATTAATGATGATGATGATGATGATGAT 1446
Db 533 LeuGlnPheThrTyrAspHisIleGlnAlaHisValAsnAspMetLeuGlyArgIleAla 552
QY 1447 ACAGCTTGTCGGATTCAGAAATAGAGAACTGTTTATGCGCAGCAGGAGGATTAAGATT 1506
Db 553 AlaAlaTyrCysGluLeuGlnAsnLysAspArgThrLeuTrpSerGluMetSerArgLeu 572
QY 1507 AATCTAGCGCTACAGCAGTCAACATTAGGAAGGAGAGTGGCTGCAAGATGTTGGGG 1566
Db 573 AsnProSerAlaValAlaThrAlaAlaLeuGlyGlnArgValCysAlaArgMetLeuGly 592
QY 1567 GATGTCGCTGCTGATCGAGCTGCACTGCTATAGATGCGGAATCGCTCAGTCTGCAAAAT 1626
Db 593 AspValMetAlaIleSerArgCysValGluValAlaArgGly---GlyValTyrValGlnAsn 611
QY 1627 TCTATGCGAGTTATCACATCCACTAATACATGTTATAGCCGACCATTTGGTTCTATTTC 1686
Db 612 SerMetArgValProGlyGluArgGlyThrCysTyrSerArgProLeuValThrPheGlu 631
QY 1687 TATGAGAAAACCAAGGAAACATACAGGACAACTCGGTGAAACCAACAGCTGCTTCCA 1746
Db 632 His-----AsnGlyThrGlyValIleGlyGlnLeuGlyAspAsnGlnLeuLeuIle 650
QY 1747 ACGCTAGAGCTGTAGAGCCATGCTCGGCTAATCATCTAGATATTTCTGTTTGGATCC 1806
Db 651 SerArgAspLeuIleGluProCysThrGlyAsnHisArgArgTyrPheLysLeuGlySer 670
QY 1807 GGTATGCTTTATTTGAACATATAATTTTGTAGATGGTAGCGCTGCCGATATACAG 1866
Db 671 GlyTyrValTyrTyrGluAspTyrAsnTyrValArgMetValGluValPro-----Glu 688
QY 1867 ATTGCTAGCATTTGCTGAGCTTAATCTACCTGCTAGAGATCGGGAATTTTGCCT 1926
Db 689 ThrIleSerThrArgValThrLeuAsnLeuThrLeuLeuGluAspArgGluPheLeuPro 708
QY 1927 TTATCCGTTTACACAAAAGAGAGTTCGCTGATGTTGCTGATGTTGCTGATGTTATGCAAGTA 1986
```

```
Db 709 LeuGluValTyrThrArgGluGluLeuAlaAspThrGlyLeuLeuAspTyrSerGluIle 728
QY 1987 GCTCGCCGCAATCAATCATACATTAATTTTATGACATAAACAAGTAAATAGAGTG 2046
Db 729 GlnArgArgAsnGlnLeuHisAlaLeuLysPheTyrAspIleAspArgValValLysVal 748
QY 2047 GATACAAAT-----TACGCGGGCTGCAGGAATTC 2076
Db 749 AspHisAsnValValLeuLeuArgGlyIleAlaAsnPhePheGlnGlyLeuGlyAspVal 768
QY 2077 GCTGCT-----ATGCTCTATTACTAAAAAGATCCAAACCCAAATAAT 2118
Db 769 GlyAlaAlaValGlyLysValValLeuGlyAlaThrGlyAlaValIleSerAlaValGly 788
QY 2119 GGCCAAACCCCAATTAGAAGCAGCGCAATGGAGTTAACAGATCTAATCAATGCTAAAGCG 2178
Db 789 GlyMetValSerPhe-----LeuSerAsnProPheGlyAlaLeuAla 802
QY 2179 ATGACATTTAGCTTCACTACAAGACTATGCCAAGATTTGAAGCTAGTGTATCATCTGCTTAT 2238
Db 803 IleGlyLeuLeuValLeu-----AlaGlyLeuValAlaAlaPhe---LeuAlaTyr 818
QY 2239 AGTGAAGCTGAACAGCTTAACAATTAACCTTAATGCAACATTAGAACAACATAAAATGGCT 2298
Db 819 ArgHisIleSerArgLeuArgArgAsnProMetLysAlaLeuTyrProVal---ThrThr 837
QY 2299 AAAACTAATTTAGAATTCAGCCATCAACCAAGCTAATACGATAAACAGACTTTTGTGTAAT 2358
Db 838 LysThrLeuLysGluAspGlyValAspGluGlyAspValAspGluAlaLysLeuAspGln 857
QY 2359 GAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACACCT 2415
Db 858 AlaArgAspMetIle-----ArgTyrMetSerIleValSerAlaLeuGluGlnGln 874
```

Search completed: October 8, 2003, 18:21:31

Job time : 133.342 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 8, 2003, 17:32:42 ; Search time 183.496 Seconds
(without alignments)
5641.626 Million cell updates/sec

Title: US-09-147-052-3

Perfect score: 5665

Sequence: 1 atgcacatttttagcgagaa.....gacccgggtacatttttataa 3261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlpl
-Q/cgn2_1/USPTO_spol/US09147052/runat_08102003_154340_29792/app_query.fasta_1.5980
-DB-A_Geneseq_19Jun03 -OPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09147052.ecgn_1.1.474/runat_08102003_154340_29792 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5608	99.0	1086	18	AAW36051	Hybrid Marek's dis
2	3601	63.6	865	14	AAW36051	Marek's Disease VI
3	2015	35.6	456	18	AAW36050	Hybrid Marek's dis
4	1968	34.7	879	16	AAW77024	Canine herpesvirus
5	1913	33.8	615	15	AAW63230	Mycoplasma gallise
6	1859	32.8	610	15	AAW63229	Mycoplasma gallise
7	1779	31.4	868	13	AAW22615	Varicella-zoster v
8	1779	31.4	913	19	AAW68404	Aujeszky's disease
9	1765.5	31.2	845	6	AAW50035	N-terminal sequenc
10	1746	30.8	933	13	AAW27807	Bovine herpes viru
11	1746	30.8	933	13	AAW41343	Bovine herpesvirus
12	1743	30.8	943	21	AAV32470	DNA encoding felin
13	1743	30.8	943	22	AAW50113	Feline herpesvirus
14	1739.5	30.7	891	17	AAW92746	B virus gb glycopr
15	1739.5	30.7	891	19	AAW70293	Simian herpesvirus
16	1736	30.6	904	23	AAE17812	Herpes simplex vir
17	1735	30.6	854	19	AAW72113	HSV-2 strain SB5 C
18	1735	30.6	904	14	AAW41778	Glycoprotein B (gB
19	1735	30.6	904	17	AAW00376	HSV-2 glycoprotein
20	1735	30.6	904	18	AAW34552	Herpes simplex vir
21	1735	30.6	904	19	AAW72193	HSV-2 strain SB5 C
22	1735	30.6	904	22	AAW74442	Herpes simplex vir
23	1732	30.6	854	8	AAW70347	Varicella-zoster v
24	1726	30.5	904	9	AAW80914	Sequence of Herpes
25	1720	30.4	907	8	AAW71136	Herpes Simplex Vir
26	1719.5	30.4	885	17	AAW92747	SAB virus gb glyco
27	1705.5	30.1	904	17	AAW00375	HSV-1 glycoprotein
28	1705.5	30.1	904	22	AAW74441	Herpes simplex vir
29	1704.5	30.1	795	19	AAW72062	HSV-2 strain SB5 C
30	1703.5	30.1	903	8	AAW70426	Recombinant herpes
31	1702.5	30.1	904	12	AAW14665	HSVgb polypeptide.
32	1700.5	30.0	903	7	AAW60244	Herpes simplex vir
33	1699.5	30.0	903	8	AAW71135	Herpes simplex vir
34	1696.5	29.9	904	14	AAW41779	Glycoprotein B (gB
35	1696.5	29.9	905	9	AAW80915	Sequence of Herpes
36	1694.5	29.9	973	8	AAW70769	Glycoprotein B Of
37	1693	29.9	973	12	AAW14680	HSV surface antige
38	1690.5	29.8	904	18	AAW34553	Herpes simplex vir
39	1682	29.7	694	12	AAW14666	Truncated HSVgb po
40	1666	29.4	368	14	AAW44493	Mycoplasma gallise
41	1666	29.4	368	15	AAW63227	Mycoplasma gallise
42	1658	29.3	368	16	AAW75955	Mycoplasma gallise
43	1601	28.3	928	16	AAW77399	BHV1 gi glycoprote
44	1576	27.8	903	6	AAW50312	Herpes simplex vir
45	1362.5	24.1	932	19	AAW44947	Bovine herpesvirus

ALIGNMENTS

RESULT 1	
AAW36051	
ID	AAW36051 standard; Protein; 1086 AA.
XX	
AC	AAW36051;
XX	
DT	15-JUL-1998 (first entry)
XX	
DE	Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX	
KW	Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW	antigen; vaccine; poultry.
XX	
OS	Chimeric - Marek's disease gammaherpesvirus.
OS	Chimeric - Mycoplasma gallisepticum.
XX	
FT	Key Location/Qualifiers
FT	Region 1..672

FT Region /note- "derived from Marek's disease virus gB protein"
 FT 693..1086
 PN /note- "derived from M. gallisepticum antigen"
 XX WO9736924-A1.
 XX 09-OCT-1997.
 XX 28-MAR-1997; 97WO-JP01084.
 XX 29-MAR-1996; 96JP-0103548.
 XX (JAPG) NIPPON ZEON KK.
 PA Saito S, Tsuzaki Y, Yanagida N;
 PI WPI; 1997-503046/46.
 DR N-PSDB; AAT96596.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;

 Alignment Scores:
 Pred. No.: 0 Length: 1086
 Score: 5608.00 Matches: 1083
 Percent Similarity: 99.82% Conservative: 1
 Best Local Similarity: 99.72% Mismatches: 2
 Query Match: 98.99% Indels: 0
 DB: 18 Gaps: 0

 US-09-147-052-3 (1-3261) x AAW36051 (1-1086)

 QY 1 ATGCACATNTTTAGGCGGAATTCATATTTTCCCTATATAGTTATCTATATGATGACGAC 60
 DB 1 MetHisTyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20

 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTTGTTTCGAGCGTCCAGTTGTCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40

 QY 121 GAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGGTTTCAACCGTGTATCCGCTA 180
 DB 41 GluGluGluSerThrPheTyrLeuCysProProValGlySerThrValIleArgLeu 60

 QY 181 GRACCCCGGAAATGTCGGAACCTAGAAAGCCAGTGGGTGGGTGAAGAAATCGCG 240
 DB 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80

 QY 241 ATATATTTAAAGAGAATATCAGTCCATATAATTTAAAGTGCACGCTTTATATATAAAT 300
 DB 81 IleLeuPheLysGlnAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100

 QY 301 ATCATTCAGACGACATGGAGGGGACATATAGACAGATCACTAATCATATACATA 360
 DB 101 IleIleGlnThrThrTrpThrGlyThrThrTyrArgGlnIleThrAsnArgTyrThr 120

 QY 361 GATAGACCCGCTTCCATTGGAAGATCAGGATCTAATCGACGGAAAGAAATGC 420
 DB 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140

 QY 421 TCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTCACAGGGATCG 480

Db 141 SerSerLysAlaArgTyrLeuArgAsnAsnValTyrValGluAlaPheAspArgAspAla 160
 QY 481 GGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACAGCCGCCGATAGGGCATGG 540
 DB 161 GlyGluLysGlnValLeuLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
 QY 541 CACACGACTAATGAGACGATACCGTGTGGGATCACCATGGATATATATCCACGGGAACC 600
 DB 181 HisThrThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
 QY 601 TCCGTCAATTGTATAGTAGAGAAATGGATGCCGCTCTGTTCGTTCCGTATTCAATTTT 660
 DB 201 SerValAsnCysIleValGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
 QY 661 GCAATGCCCAATGCGACATCGCGAATATCCATTTTATGTTATGTTATCCACACAGAG 720
 DB 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProGlu 240
 QY 721 GCTCCGCGAGAACCCATGGGATATCCCGAGTAATTTCAAACAATAGATAGCTATTTT 780
 DB 241 AlaAlaGluProMetGlyTyrProGlnAspAsnPheLysGlnLeuAspSerTyrPhe 260
 QY 781 TCAATGGATTGGACAAGCGTCGAAAAGCAAGCCTTCCAGTCAAGCGTAATTTCTCATC 840
 DB 261 SerMetAspLeuAspLysArgLysAlaSerLeuProValLysArgAsnPheLeuIle 280
 QY 841 ACATCACATTCACAGTGGTGGCTGGCTGCCAAACTACTCGTGTATCTCAATG 900
 DB 281 ThrSerHisPheThrValGlyTrpAspTrpAlaProLysThrThrArgValCysSerMet 300
 QY 901 ACTAAGTGGAAAGAGGTGACTGAAATGTTCCGTGCAACAGTTAATGGAGATACAGATTT 960
 DB 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
 QY 961 ATGCGCGTGAACCTTCGGCAAGCTTTATCAGTAATACGACTGAGTTGTATCCAAATCGC 1020
 DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
 QY 1021 ATCATATTAGCAATGTATTAAACGCGAGGAGAGACAGCAATCGAGCAGATATTAGG 1080
 DB 341 IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
 QY 1081 ACAAATATAATGACAGTCACGTCACAGTTGGACATGTACAATATTTCTTGGCTCGGG 1140
 DB 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
 QY 1141 GGATTTATTGTAGCATATCAGCCTGTCTATCCAAATCCCTGGCTCATATGTACCTCAGA 1200
 DB 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
 QY 1201 GAATTGATGAGAGACAACAGGACCGATGAGATGCTCGACTGCTAAACAATAAGCATGCA 1260
 DB 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
 QY 1261 ATTTATAAGAAAATGCTACCTCATGTCCAGTTCGCGGAGATATTCGAAATGACCA 1320
 DB 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgArgAspIleArgAsnAlaPro 440
 QY 1321 AATAGAAAATACATTTAGACGACACAGCTATTAAATCGACATCGCTGTCTCAATTC 1380
 DB 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerSerValGlnPhe 460
 QY 1381 GCCATCTCCAATTTCTTTATGATCATATACAACCCATATTAATGATATGTTTAGTAGG 1440
 DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
 QY 1441 ATTGCCACAGTGGTGGGAATTCGAAATAGAGAATTTGTTTATGGCACGAAGGATA 1500
 DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGluGlyIle 500
 QY 1501 AAGATTAACTTCAGCCCTACGCGAGTGCACACATTAGGAGGAGAGTGGCTGCAAGATC 1560

Db 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgValAlaAlaLysMet 520
QY 1561 TTGGGGGATCGCTGCTGATCGAGCTGCCTCTATAGATCGGAATCCGTCACCTTTG 1620
Db 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
QY 1621 CAAATCTATGCGAGTTATCACATCCACTCAATACATGTTATAGCCGACCAATGGTCTA 1680
Db 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560
QY 1681 TTTTTCATATGAGAGAAACCAAGGAAACATACAGGACAACTCGGTGAAACCAACGAGTTG 1740
Db 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyLeuGlyGluAsnAsnGluLeu 580
QY 1741 CTTCCACGCTAGAGCGTGTAGACCAATGCTCGGCTCAATCATCGTAGATATTTCTGTTT 1800
Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe 600
QY 1801 GGATCCGGTTATGCTTTATTTGAAACTATATTTCTTAAGATGGTAGCGCTGCCGAT 1860
Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAlaAsp 620
QY 1861 ATACAGATTTGCTAGCACATTTGCTGAGCTTAATCTAACCTTCGCTAGAGATCGGGAATTT 1920
Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuLeuGlnAspArgGluIle 640
QY 1921 TTGCTTTATCCGTTTACACAAAAGAGAGTTGCGTGATGTTGGTGATTTGGATTATGCA 1980
Db 641 LeuProLeuSerValTyrThrLysGluLeuArgAspValGlyValPheAspTyrAla 660
QY 1981 GAATGAGTCCCGCACTCACTACATCACTTAATTTTATGATGACATAAACAAGTANTA 2040
Db 661 GluValAlaArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
QY 2041 GAATGGATACAAATTTACGCGGGCTGCAGGAATTTGCGCTGATGCTTACTTACTAAAAA 2100
Db 681 GluValAspThrAsnTyrAlaGlyLeuGlnGluPheGlyCysMetSerIleThrLysLys 700
QY 2101 GATGCAAAACCAATTAATGGCCAAACCAATTAGAAGCAGCGCGGAATGAGTTAACAGAT 2160
Db 701 AspAlaAsnProAsnAsnGlyGlnThrGlnLeuGluAlaAlaArgMetGluLeuThrAsp 720
QY 2161 CTAATCAATGCTAAAGCGATGACATTAAGTCTTCACTACAGACTATGCCAAGATTGAAGCT 2220
Db 721 LeuIleAsnAlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAla 740
QY 2221 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTTAAACAATAACCTTAATGCAACATTA 2280
Db 741 SerLeuSerSerAlaTyrSerGluAlaGluThrValAsnAsnAsnLeuAsnAlaThrLeu 760
QY 2281 GAACAACTAAATGGCTGCTAAACTAATTTAGATCAGCCATCAACCAAGCTTAATACGGAT 2340
Db 761 GluGlnLeuLysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsnThrAsp 780
QY 2341 AAAAGCAGCTTTTGATAATGAACCCCAATTTAGTTGAAGCATACAAAGCACTAAACACC 2400
Db 781 LysThrThrPheAspAsnGlnHisProAsnLeuValGluAlaTyrLysAlaLeuLysThr 800
QY 2401 ACTTTAGAACACGCTGCTACTAACTTGAAGCTTTGTCATCAACTGCTTATAATCAAAATTT 2460
Db 801 ThrLeuGluGlnArgAlaThrAsnLeuGluGlyLeuSerSerThrAlaTyrAsnGlnIle 820
QY 2461 CGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAACTAACTAAACACACTAGAT 2520
Db 821 ArgAsnAsnLeuValAspLeuTyrAsnLysAlaSerSerLeuIleThrLysThrLeuAsp 840
QY 2521 CCACATAATTTAGGGAGCCTTTTATGATTCCTAATGAGATTACTACAGCTAATAAGAAATTT 2580
Db 841 ProLeuAsnGlyGlyThrLeuLeuAspSerAsnGluIleThrThrAlaAsnLysAsnIle 860
QY 2581 AATAATAGCTTATCAACTATTAAATCAACAAAGACTAATGCTGATGCTTATCTAATAGT 2640
Db 861 AsnAsnThrLeuSerThrIleAsnGlnLysThrAsnAlaAspAlaLeuSerAsnSer 880

QY 2641 TTTTATTAAGAGTATTCAAAAATAATGAACAAAAGTTTGTAGGACTTTTACAAACGCT 2700
Db 881 PheIleLysLysValIleGlnAsnAsnGluGlnGlyPheValGlyThrPheThrAsnAla 900
QY 2701 AATGTTCAACCTTCAAACTACAGTTTGTCTTCTTTAGTGTGATGTAAACCCGCTCAAT 2760
Db 901 AsnValGlnProSerAsnTyrSerPheValAlaPheSerAlaAspValThrProValAsn 920
QY 2761 TATAAATATGAAGAAGACCGTTTGGAAATGGTGATGAACCTTCAAGTAGAATTCCTGCA 2820
Db 921 TyrLysTyrAlaArgArgThrValTrpAsnGlyAspGluProSerSerArgIleLeuAla 940
QY 2821 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAAACAACACGAG 2880
Db 941 AsnThrAsnSerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsnThrLys 960
QY 2881 TACCAATTTAGCTTTTAGCAACTATGGTCCATCACTGCTTATTTATATTTCCCTTATAAG 2940
Db 961 TyrGlnPheSerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLys 980
QY 2941 TTGTTAAAGCAGCTGATGCTTAATAACCTTGGATTACAAATACAAATTAATAATGGAAT 3000
Db 981 LeuValLysAlaAlaAspAlaAsnAsnValGlyLeuGlnTyrLysLeuAsnAsnGlyAsn 1000
QY 3001 GTTCAACAAGTTGAGTTTGGCCACTTCAACTAGTGCATAATAATACACTACAGCTTAATCCAAT 3060
Db 1001 ValGlnGlnValGluPheAlaThrSerThrSerAlaAsnAsnThrThrAlaAsnProThr 1020
QY 3061 CCACAGTTGATGAGATTAAAGTTGCTAAAGTTCCTAAATCGTTTATCAGGTTTAAGATTGGCCAA 3120
Db 1021 ProAlaValAspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPheGlyGln 1040
QY 3121 AACCAATCGAATTAAGTGTTCACACGGGTGAAGGAAATATGAATAAAGTTGCGCCAATG 3180
Db 1041 AsnThrIleGluLeuSerValProThrGlyGluGlyAsnMetAsnLysValAlaPromet 1060
QY 3181 ATTGGCAACATTTATCTTAGCTCAATGAAATAATGCTGATAAGATCCCGGGTACCCT 3240
Db 1061 IleGlyAsnIleTyrLeuSerSerAsnGluAsnAsnAlaAspLysIleProGlyTyrArg 1080
QY 3241 CGACCCGGTACATTTTAA 3258
Db 1081 ArgProGlyThrPheLeu 1086
RESULT 2
AAR30169
ID AAR30169 standard; Protein; 865 AA.
XX
AC AAR30169;
XX
DT 25-MAR-2003 (updated)
DT 07-MAY-1993 (first entry)
XX
DE Marek's Disease Virus glycoprotein B homologue of HSV.
XX
KW Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
KW Herpes Simplex Virus.
XX
OS Marek's Disease Virus.
XX
PN EP520753-Al.
XX
PD 30-DEC-1992.
XX
PF 24-JUN-1992; 92EP-0305775.
XX
PR 28-JUN-1991; 91US-0722860.
PR 10-DEC-1991; 91US-0803633.
XX
PA (JAPG) NIPPON ZEON KK.
PA (USDA) US SEC OF AGRIC.
XX

PI Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
 XX WPI; 1993-001546/01.
 XX Recombinant fowl pox virus contg. Marek's disease virus antigen
 PT gene - used to produce cell-free vaccine against Marek's disease
 PT virus
 XX
 PS Example 2; Page 15-19; 30pp; English.
 XX
 CC The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
 CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-Sali
 CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
 CC from K3 fragment were ligated with EcoRI, Sali digested pUC18. The
 CC sequence of the putative MDV gBh was determined by sequencing a set
 CC of deletion mutants. The nucleotide and amino acid sequences were
 CC found to be identical with the published sequences of the gBh of
 CC RBIB strain of MDV (Ross et al., J. Gen. Virol., 70:1789-1894, 1988).
 CC A fragment contg. the entire coding region of MDV gBh was inserted
 CC into pNZ1729R (see AQ034774-Q34778) to produce a recombinant FPV/MDVgBh
 CC virus for immunising chickens. See also AQ034780-Q34781.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 865 AA;
 Alignment Scores:
 Pred. No.: 6, 28e-275 Length: 865
 Score: 3601.00 Matches: 719
 Percent Similarity: 87.81% Conservative: 16
 Best Local Similarity: 85.90% Mismatches: 60
 Query Match: 63.57% Indels: 42
 DB: 14 Gaps: 8
 US-09-147-052-3 (1-3261) x AAR30169 (1-865)
 QY 1 ATGCACATATTTAGCGGGAATTCATATTTTCCCTATATAGTTATCTATATGTCACGAAC 60
 DB 1 MethistyrPheArgArgAsnCysIlePhePheLeuIleValIleLeuTyrGlyThrAsn 20
 QY 61 TCATCTCCGAGTACCCAAATGTGACATCAAGAGAAGTGTTCGAGCGTCCAGTGTCT 120
 DB 21 SerSerProSerThrGlnAsnValThrSerArgGluValValSerSerValGlnLeuSer 40
 QY 121 GAGGAAGAGTCTAGCTTTATCTTTTTCCTCCCGCCAGTGGTGTCAACCGGTATCCGCTA 180
 DB 41 GluGluGluSerThrPheThrLeuLeuValValSerProValGlySerThrValIleArgLeu 60
 QY 181 GAACCGCGCGAAATGTCGGAACCTAGAAAGCCACCGAGTGGGTGAGGAATCGCG 240
 DB 61 GluProProArgLysCysProGluProArgLysAlaThrGluTrpGlyGluGlyIleAla 80
 QY 241 ATATTATTAAGAGATATCAGTCCATATAAATTTAAAGTGAGCGCTTTATTAATAAAT 300
 DB 81 IleLeuPheLysGluAsnIleSerProTyrLysPheLysValThrLeuTyrTyrLysAsn 100
 QY 301 ATCATTCAGACGACATGACGGGACGACATATATAGAGATCACTAATTCGATATACA 360
 DB 101 IleIleGlnThrThrTrpThrGlyThrThrArgGlnIleThrAsnArgTyrThr 120
 QY 361 GATAGGACCGCGTTTCATTAAGAGATCAAGATCTAATTCGACGCGCAAGGAATGCG 420
 DB 121 AspArgThrProValSerIleGluIleThrAspLeuIleAspGlyLysGlyArgCys 140
 QY 421 TCATCTAAAGCAAGATACCTTAGAACAATGTATATGTTGAAGCGTTTGACAGGGATGCG 480
 DB 141 SerSerLysAlaArgTyrLeuArgAsnValValGluAlaPheAspArgAspAla 160
 QY 481 GGAGAAAACAAGTACTTCTAAACCATCAAAATTCACACCGCCCGAATCTAGGCGATGG 540
 DB 161 GlyGluLysGlnValLeuLysProSerLysPheAsnThrProGluSerArgAlaTrp 180
 QY 541 CACAGCACTAATGACAGCTATACCGTGTGGGATCACCATCGATATATCGAACCGGAAC 600

DB 181 HistThrAsnGluThrTyrThrValTrpGlySerProTrpIleTyrArgThrGlyThr 200
 QY 601 TCCGTCATTTGTATAGTAGAGAAATGATGCCGCTCTGTGTTCGCTATTCATATTTT 660
 DB 201 SerValAsnCysIleValGluGluMetAspAlaArgSerValPheProTyrSerTyrPhe 220
 QY 661 GCAATGGCCAATGGCGACATCGGAACATATCTCCATTTTATGGTCTATCCCAACGAG 720
 DB 221 AlaMetAlaAsnGlyAspIleAlaAsnIleSerProPheTyrGlyLeuSerProProGlu 240
 QY 721 GCTCGCGCAGAACCCATGGGATATCCCGAGATATTTTCAACCACTAGATAGTATTTT 780
 DB 241 AlaAlaAlaGluProMetGlyTyrProGlnAspAsnPhelLysGlnLeuAspSerTyrPhe 260
 QY 781 TCAATGGATTTGGACAAGCGTCCGAAAGCAAGCCCTCCAGTCAAGCGTAACTTTCTCATC 840
 DB 261 SerMetAspLeuAspLysArgArgLysAlaSerLeuProValLysArgAsnPhelLeu 280
 QY 841 ACATCACACTTTCACAGTGGGTGGGAGTGGGCTCCAAAACACTACTGCTGTATGTTCAATG 900
 DB 281 ThrSerHisPheThrValGlyTyrAspTrpAlaProLysThrThrArgValCysSerMet 300
 QY 901 ACTAAGTGGAAAGAGTGAATGTTGCGTGCACAGTAAATGGGAGATACAGATT 960
 DB 301 ThrLysTrpLysGluValThrGluMetLeuArgAlaThrValAsnGlyArgTyrArgPhe 320
 QY 961 ATGGCCCGTAACTTTCCGCAACGTTTATCAGTATACAGTGAAGTATTCCTGCTCTCGGG 1020
 DB 321 MetAlaArgGluLeuSerAlaThrPheIleSerAsnThrThrGluPheAspProAsnArg 340
 QY 1021 ATCATATPAGCAATGTATTAACCGGAGGAGAGAGCAGCAATCGACAGATATTTAGG 1080
 DB 341 IleIleLeuGlyGlnCysIleLysArgGluAlaGluAlaIleGluGlnIlePheArg 360
 QY 1081 ACAAAATATAATGACAGTCACGTCAGGTGGACATGTACAATATTTCTTGCTCTCGGG 1140
 DB 361 ThrLysTyrAsnAspSerHisValLysValGlyHisValGlnTyrPheLeuAlaLeuGly 380
 QY 1141 GGATTTATGTAGCATATCAGCTGTCTTATCCAAATCCCTGGCTCATATGACCTCAGA 1200
 DB 381 GlyPheIleValAlaTyrGlnProValLeuSerLysSerLeuAlaHisMetTyrLeuArg 400
 QY 1201 GAATTCATGAGACACACAGGACCGATGAGATGCTCGACCTGTAACCAATTAAGATGCA 1260
 DB 401 GluLeuMetArgAspAsnArgThrAspGluMetLeuAspLeuValAsnAsnLysHisAla 420
 QY 1261 ATTATTAAGAAAATGCTACCTCATTTGTCACGATTCGCGGAGATATTCGAAATGCACCA 1320
 DB 421 IleTyrLysLysAsnAlaThrSerLeuSerArgLeuArgAspIleArgAsnAlaPro 440
 QY 1321 AATAGAAAATACATTAGACACACACAGCTATTAATTCGACATCGTCTGTTCAATTC 1380
 DB 441 AsnArgLysIleThrLeuAspAspThrThrAlaIleLysSerThrSerValGlnPhe 460
 QY 1381 GCCATGCTCAATTTCTTTATGATCATATACAACCCATTAATGATATGTTAGTAGG 1440
 DB 461 AlaMetLeuGlnPheLeuTyrAspHisIleGlnThrHisIleAsnAspMetPheSerArg 480
 QY 1441 ATGGCCACAGCTTGGTGGCAATTCGAGAATAGAGAATTTGTTTTATGGCACGAAGGATA 1500
 DB 481 IleAlaThrAlaTrpCysGluLeuGlnAsnArgGluLeuValLeuTrpHisGlyIle 500
 QY 1501 AAGATTAATCCTAGCGCTACAGGAGTGCACATTAGGAGGAGAGTGGCTGCAAGATG 1560
 DB 501 LysIleAsnProSerAlaThrAlaSerAlaThrLeuGlyArgArgValAlaAlaLysMet 520
 QY 1561 TTGGGGATGCTGCTGCTGTATCGAGTGCACCTGCTATAGATCGGAATCCGTCACCTTTG 1620
 DB 521 LeuGlyAspValAlaAlaValSerSerCysThrAlaIleAspAlaGluSerValThrLeu 540
 QY 1621 CAAATTTCTATGGAGTTATCACATCCACCTAATACATGTTATAGCCGACCATTTGGTCTA 1680
 DB 541 GlnAsnSerMetArgValIleThrSerThrAsnThrCysTyrSerArgProLeuValLeu 560

QY 1681 TTTTCATATGAGAAACCAAGCAACATACAGGCAACTCGGTGAAACACAGGCTTG 1740
 Db 561 PheSerTyrGlyGluAsnGlnGlyAsnIleGlnGlyGluAsnAsnGluLeu 580
 QY 1741 CTTCCACGCTAGAGCGCTAGAGCCATGCTCGGCTAATCATCATCTAGATATTTCTCTTT 1800
 Db 581 LeuProThrLeuGluAlaValGluProCysSerAlaAsnHisArgTyrPheLeuPhe 600
 QY 1801 GGATCCGGTTATGCTTTATTGAAACTATAATTTTGAAGATGGTAGCGCTGCCGAT 1860
 Db 601 GlySerGlyTyrAlaLeuPheGluAsnTyrAsnPheValLysMetValAspAlaAsp 620
 QY 1861 ATACAGATTGCTAGCACATTGCTCGACCTTAATCTAACCTGCTAGAAATCGGAAATT 1920
 Db 621 IleGlnIleAlaSerThrPheValGluLeuAsnLeuThrLeuGluAspArgGluIle 640
 QY 1921 TTGGCTTTATCCGTTTACAAAAGAGAGTGGCTGATGTTGGTGATTGGATTATGCA 1980
 Db 641 LeuProLeuSerValTyrThrLysGluGluLeuArgAspValGlyValLeuAspTyrAla 660
 QY 1981 GAAGTAGCTCGCGCAATCACTACATGAATTAATTTATGACATAAAACAAGTAATA 2040
 Db 661 GluValAlaArgArgAsnGlnLeuHisGluLeuLysPheTyrAspIleAsnLysValIle 680
 QY 2041 GAAGTGATACAAATTAACGGC-----GGCTGCAGGAATTCGGCTGTATGCTATT 2091
 Db 681 GluValAspThrAsnTyrAlaPheMetAsnGlyLeuAlaGluLeu----- 695
 QY 2092 ACTAAAAAGATGCAACCCCAATAATGCGCAACCCCAATA---GAAGCAGCGCGAATG 2148
 Db 696 -----PheAsnGlyMetGlyGlnValGlyGlnAlaIleGlyLys 708
 QY 2149 GAGTTAAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACATACAGACTATGCC 2208
 Db 709 ValValValGlyAlaAlaGlyAlaIleValSerThrIleSerGlyValSerAlaPheMet 728
 QY 2209 AAG-----ATTTGAAGCTAGTTTATCATCTGCT 2235
 Db 729 SerAsnProPheGlyAlaLeuAlaIleGlyLeuIleIleAlaGlyLeuValAlaAla 748
 QY 2236 TATAGTGAAGCTGAACAGTTAAC-----AATAACCTTAATGCA----- 2274
 Db 749 PheLeuAlaTyrArgTyrValAsnLysLeuLysSerAsnProMetLysAlaLeuTyrPro 768
 QY 2275 ---ACATTAGAACAACTAAAATAGCTAAACAACTAATTTAGAAATCAGCCATCAACCAAGCT 2331
 Db 769 MetThrThrGluValLeuLysAlaGlnAlaThrArgGluLeuHisGlyGluSerAsp 788
 QY 2332 AATACGGATAAACGACTTTTGTATAATGACACCCCAATTTAGTTGAAGCA----- 2382
 Db 789 AspLeuGluArgThrSerIleAspGluArg-----LysLeuGluGluAlaArgGluMet 806
 QY 2383 -----TACAAAGCAGCTAAAACCCACTTTAGAACACAGCTGCTACTACCTT 2427
 Db 807 IleLysTyrMetAlaLeuValSerAlaGluGluArgHisGluLysLysLeu 823
 RESULT 3
 AAW36050
 ID AAW36050 standard; Protein; 456 AA.
 XX
 AC AAW36050;
 XX
 DT 15-JUL-1998 (first entry)
 XX
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.
 XX
 KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 XX
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX

EH Key Location/Qualifiers
 FT Region 1..64
 FT /note= "derived from Marek's disease virus gb protein"
 FT Region 65..456
 FT /note= "derived from M. gallisepticum antigenic protein"
 XX WO9736924-A1.
 PN 09-OCT-1997.
 XX 28-MAR-1997; 97WO-JP01084.
 XX 29-MAR-1996; 96JP-0103548.
 XX (JAPG) NIPPON ZEON KK.
 XX Saito S, Tsuzaki Y, Yanagida N;
 DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; Sipp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gb fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 456 AA;
 Alignment Scores:
 Pred. NO.: 4,05e-150 Length: 456
 Score: 2015.00 Matches: 396
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 35.57% Indels: 0
 DB: 18 Gaps: 0
 US-09-147-052-3 (1-3261) x AAW36050 (1-456)
 QY 2071 GAATTCGCTGTATGCTCTATTACTAAAGATGCAACCAACCAATATGGCCAAACCCAA 2130
 Db 61 GluPheGlyCysMetSerIleThrLysLysAspAlaAsnProAsnAsnGlyGlnThrGln 80
 QY 2131 TTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCT 2190
 Db 81 LeuGluAlaAlaArgMetGluLeuThrAspLeuIleAsnAlaLysAlaMetThrLeuAla 100
 QY 2191 TCACCTAAGACTATGCCAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAA 2250
 Db 101 SerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSerSerAlaTyrSerGluAlaGlu 120
 QY 2251 ACAGTTAACATAACCTTAATGCAACATTTAGAACCACTAAATGCTGCTAACTAATTTA 2310
 Db 121 ThrValAsnAsnLeuAsnAlaThrLeuGluGlnLeuLysMetAlaLysThrAsnLeu 140
 QY 2311 GAATCAGCCCATCAACCAAGCTAATACGATATAAACAGACTTTTGTGTAATGAACACCAAT 2370
 Db 141 GluSerAlaIleAsnGlnAlaAsnThrAspLysThrThrPheAspAsnGluHisProAsn 160
 QY 2371 TTAGTTGAAGCATACAAAGCACTAAAACCCACTTTAGAACCACTGCTACTACTGCTGAA 2430
 Db 161 LeuValGluAlaTyrLysAlaLeuLysThrThrLeuGluGlnArgAlaThrAsnLeuGlu 180
 QY 2431 GGTTTGTATCACTGCTTATATCAAAATTCGCAATAATTTAGTGGATCTATACAAATAA 2490
 Db 181 GlyLeuSerSerThrAlaTyrAsnGlnIleArgAsnAsnLeuValAspLeuTyrAsnLys 200

2491	GCTAGTAGTTTAATAACATAAAACACATAGATCCACTAAATGGGGGAACGCTTTTAGATTCT	2550
2501	AlaSerSerLeuIleThrLysThrLeuAspProLeuAsnGlyGlyThrLeuLeuAspSer	220
2551	AATGAGATTACTACAGCTAATAAGAATATTAAATACGTTATCAACTATTAAATGAACAA	2610
221	AsnGluIleThrThrAlaAsnLysAsnIleAsnAsnThrLeuSerThrIleAsnGluGln	240
2611	AAGACTAATAGCTGATGATTAATCTAATAGATTTTATTAAAAAAGTGATCCAAATAATGAA	2670
241	LysThrAsnAlaAspAlaLeuSerAsnSerPheIleLysLysValIleGlnAsnAsnGlu	260
2671	CAAAAGTTTGTAGGAGCTTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGTTTGT	2730
261	GlnSerPheValGlyThrPheThrAsnAlaAsnValGlnProSerAsnTyrSerPheVal	280
2731	GCTTTTAGTCGTGATGAACACCCGCTCAATTATAATATCATCAGAGAGGCCGTTGGAA	2790
281	AlaPheSerAlaAspValThrProValAsnTyrLysTyrAlaAlaArgThrValThrAsn	300
2791	GGTGATGAACCTTCAAGTAGAATCTTGTCAAAACACGAATAGTATCATCAGATGTTCTGG	2850
301	GlyAspGluProSerSerArgIleLeuAlaAsnThrAsnSerIleThrAspValSerTrp	320
2851	ATTATAGTTTAGCTGGGAACAAACACGAAGTACCAANTTAGTTTAGTCAACTATGTC	2910
321	IleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPheSerPheSerAsnTyrGlyPro	340
2911	TCAACTGGTTATTATTATTTCCCTTATTAAGTTGGTTAAAGCAGCTGATGCTAATAACG	2970
341	SerThrGlyTyrLeuTyrPheProTyrLysLeuValLysAlaAlaAspAlaAsnAsnVal	360
2971	GGATTACAATACAAATTAATTAATGGAATGTTCAACAAGTTGAGTTTGGCACTTCAACT	3030
361	GlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnValGlnValGluPheAlaThrSerThr	380
3031	AGTCAAAATAACTACTACACTAATCCAACTCAACAGCTGATGAGATTAAGTTGCTAAA	3090
381	SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys	400
3091	ATCGTTTTATCAGGTTTAAGATTGGCCAAAACAACTCGAATTAAAGTTGCCACCGGT	3150
401	IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly	420
3151	GAAAGAAATATGAATAAGTTGGCGCAATGATTGGCAACTTATCTTAGCTCAAAATGAA	3210
421	GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu	440
3211	AATAATGCTGATAAGATCCCGGGTACCGTCCAGCCGGTACATTTTAA 3258	
441	AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu 456	
RESULT 4		
AAR77024		
ID	AAR77024 standard; Protein; 879 AA.	
XX	AAR77024;	
XX	20-APR-1996 (first entry)	
XX	Canine herpesvirus gB homologue.	
DE		
XX	CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;	
KW	canarypox virus; ALVAC; vCP320; dog; puppy; immunisation; antigen;	
KW	vaccine.	
XX		
OS	Canine herpesvirus.	
XX		
XX	Key	Location/Qualifiers
FH	Region	725..741
FT		/label- Transmembrane_region
FT	Region	747..771
FT		

3031	QY	AGTCAAATAA	TACTACAGCTAATCAACTCCAGCAGTTGATGAGATTAAGATGCTTAA	3090
381	Db	SerAlaAsnAsnThrThrAlaAsnProThrProAlaValAspGluIleLysValAlaLys	400	
391	QY	ATCGTTTTATCAGGTTTAAGATTGGCCAAAACACAATCGAATTAAAGTGTCCCAACGGT	3150	
401	Db	IleValLeuSerGlyLeuArgPheGlyGlnAsnThrIleGluLeuSerValProThrGly	420	
3151	QY	GAGCGAAATATGAATAAAGTTGGCCCAATGATTGGCAACATTATCTTAGCTCAAATGAA	3210	
421	Db	GluGlyAsnMetAsnLysValAlaProMetIleGlyAsnIleTyrLeuSerSerAsnGlu	440	
3211	QY	AATAATGCTGATGAAGATCCCGGGTACCGTCACCCGGTACATTTTTTA	3258	
441	Db	AsnAsnAlaAspLysIleProGlyTyrArgArgProGlyThrPheLeu	456	
RESULT 4				
AAR77024				
ID	AAR77024 standard; Protein; 879 AA.			
XX	AAR77024;			
XX	20-APR-1996 (first entry)			
XX	Canine herpesvirus gb homologue.			
DE				
XX	CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;			
KW	canarypox virus; ALVAC; vCP320; dog; puppy; immunisation; antigen;			
KW	vaccine.			
XX				
OS	Canine herpesvirus.			
XX				
Key	Location/Qualifiers			
FH	725..741			
FT	Region			
FT	/label= Transmembrane_region			
FT	Region			
FT	747..771			

Alignment Scores:			
Pred. No.:	2,53e-146	Length:	879
Score:	1968.00	Matches:	392
Percent Similarity:	62.66%	Conservative:	155
Best Local Similarity:	44.90%	Mismatches:	258
Query Match:	34.74%	Indels:	68
DB:	16	Gaps:	10

US-09-147-052-3 (1-3261) x AAR77024 (1-879)			
QY	25	ATATTTTTCCTATATGATTATCTTA	-----
Db	8	llephepellelleIythrLeuIlelleCysAspProthrThrProgluSerThrIle	
QY	49	-----TATGTATGAACACTCATCTCCGAGTAGTCCCAAAATGTGCATCATCAAGAGAA	
Db	28	AsnProLeuAsnHisHisAsnLeuSerThrProLysProthrSerAspIleargGlu	
QY	97	GTGTGTTTCGAGCGTCCAGTTG---TCTGAGGAAGAGCTACGTTTATCTTTGTGTCCTCCCA	
Db	48	IleleuargGluSerGlnIleGluSerAspAspThrSerThrPhenylMetCysProPro	
QY	154	CCAGTGGGTTCACCGTATCCGCTATAGAACCCGCCGGAATATGTCCTCGAACCTAGAAA	
Db	68	ProSerGlySerThrLeuValArgLeuGluProProArgAlaCysProAsnTyrLysLeu	
QY	214	GCCACCGAGTGGGTGAAGNAATCGCATATTTTAAAGAGAAATATCAGTCCATATAAA	
Db	88	GlyLysAsnPhThrGluGlyIleAlaValIlelelhelysGluAsnIleSerProTyrLys	
QY	274	TTTAAAGTGACCTTTATATAAAATATCATTCACAGCAGCATGTGACGGGGACGACA	
Db	108	PhelYsAlaAsnIleTyrIlyAsnIleIleIleIleThrValTrpSerGlySerThr	
QY	334	TATAGACAGTCACTTAATTCGATATACAGATAGGACGCCCTTTTCCATTGTAAGAGATCAGC	

Db	128	TyrAlaValIleThrAsnArgTyrThrAspArgValProIleGlyValProGluIleThr	147
QY	394	GATCTAATCGAGCGCAAGAGATGCTCATCTAAAGCAAGATACCTTAGAACCAATGTA	453
Db	148	GluLeuIleAspArgGlyMetCysLeuSerLysAlaAspTyrIleArgAsnAsnTyr	167
QY	454	TATGTTGAAGCGTTTGACAGGATCGGGAGAGAAACAAGTACTTCTAAACCATCAAA	513
Db	168	GluPheThrAlaPheAspLysAspGluAspProArgGluValHisLeuLysProSerLys	187
QY	514	TTCAACACGCCGAATCTAGGCGATGGCACGACTAATGAGACCTATACCGTGGGGA	573
Db	188	PheAsnThrProGlySerArgGlyTrpHisThrValAsnAspThrTyrThrLysIleGly	207
QY	574	TCACCATGGATATATCGAACGGGAACCTCCGTCAAATGTATAGTAGAGAAATGGATGCC	633
Db	208	GlySerGlyPheTyrHisSerGlyThrSerValAsnCysIleValGluGluValAspAla	227
QY	634	CGCTCTGTGTTCCGTAATTCATATTTTGCATGGCCAATGGGACATCGGNACATATCT	693
Db	228	ArgSerValTyrProTyrAspSerPheAlaIleSerThrGlyAspIleIleHisMetSer	247
QY	694	CCATTTTATGGTCTATCCACACGAGGCTGCCGAGAACCCATGGGATATCCCCAGGAT	753
Db	248	ProPhePheGlyLeuArg---AspGlyAlaHisThrGluTyrIleSerTyrSerThrAsp	266
QY	754	AATTTCAACAACACTAGATAGCTATTTTCAATGGATTTGGACAACGCTCGAAAGCAAGC	813
Db	267	ArgPheGlnGlnIleGluGlyTyrTyrProIleAspLeuAspThrArgLeuGlnLeuGly	286
QY	814	CTTCAGTCAMCGGTAACCTTCTCATCATCACACTTCACAGTTGGGTGGGACTGGGCT	873
Db	287	AlaProValSerArgAsnPheLeuThrThrGlnHisValThrValAlaTrpAsnTrpVal	306
QY	874	CAAAAACACTACGCTGTATGTTCAATGCTAAGTGGAAAGAGTGACTGAAATGTTGGCT	933
Db	307	ProLysIleArgGluValCysThrLeuAlaLysTrpArgGluIleAspGluIleIleArg	326
QY	934	GCAACAGTTAATGGGAGATACAGATTTATGGCCCGTGAACTTTCGGCAACGTTTATCAGT	993
Db	327	AspGluTyrLysGlySerTyrArgPheThrAlaLysSerIleSerAlaThrPheIleSer	346
QY	994	AATAGACTGAGTTGATCCCAATCGCATATATTAGGCATATGATTAAGCCGAGGCA	1053
Db	347	AspThrGlnPheAspIleAspArgValLysLeuSerAspCysAlaLysArgGluAla	366
QY	1054	GAAGCAGCAATCGACGATATTTAGGACAAATATAATGACAGTCAGCTCAAGTTGGA	1113
Db	367	IleGluAlaIleAspLysIleTyrLysLysTyrAsnLysThrHisIleGlnThrGly	386
QY	1114	CATGTACAAATTTCTCGCTCCGGGATTTATGTAGCATATCAGCCTGTTCTATCC	1173
Db	387	GluLeuGluThrTyrLeuAlaArgGlyGlyPheIleIleAlaPheArgProMetIleSer	406
QY	1174	AAATCCCTGGCTCATGTACCTCAGAAATTTGATGAGAGACAACAGGACCCGATGAGT	1233
Db	407	AsnGluLeuAlaLysLeuTyrIleAsnGluLeuValArgSerAsnArgThrValAspLeu	426
QY	1234	CTCGACCTGGTAAACAATAAGCATGCAATTTATAAGAAAAATGCTACCTCATTTGTCACGA	1293
Db	427	LysSerLeuLeuAsnProSerValArgGlyGlyAlaArgLysArgSerValGluGlu	446
QY	1294	TTGCGCGGAGATATTTCGAATGACCAATAAGAAAAATAACATTAGAGCACACCAAGCT	1353
Db	447	AsnLysArgSerLysArgAsnIleGluGlyIleGluAsnValAsnAsnSerThrIle	466
QY	1354	ATTAATCGACATCTCTGTTCAATTCGCCATGCTCCCAATTTCTTATGATCATATACAA	1413
Db	467	IleLysThrThrSerSerValHisPheAlaMetLeuGlnPheAlaTyrAspHisIleGln	486
QY	1414	ACCCATATTATGATATGTTAGTAGGATGCGCAGCTTGGTGGCAATTCAGAAATAGA	1473
Db	487	SerHisValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysAsnLeuGlnAsnLys	506
QY	1474	GAACTTGTTTTATGCGACGAGGATAAAGATTAATCTTAGCGCTACAGCGAGTGCAACA	1533
Db	507	GluArgThrLeuTyrAsnGluValMetLysLeuAsnProThrSerValAlaSerValAla	526
QY	1534	TTAGGAAGGAGAGTGGCTGCAAAAGATGTTGGGGGATGTCGCTGTATCGAGCTGCACCT	1593
Db	527	MetAspGlnArgValSerAlaArgMetLeuGlyAspValLeuAlaValThrGlnCysVal	546
QY	1594	GCTATAGATCGGAATCCGTCACCTTTGCAAAATTTCTATCGAGTTATCATCTCCACTAAT	1653
Db	547	AsnIleSerGlySerSerValPheIleGlnAsnSerMetArgValLeuGlySerThrThr	566
QY	1654	ACATGTTTATAGCCGACCATTTGTTCTTATTTTCATATGGAGAAACAACGAAC---ATA	1710
Db	567	ThrCysTyrSerArgProLeuIleSerPheLysAlaLeuGluAsnSerThrAsnTyrIle	586
QY	1711	CAGGACAACTCGGTGAAAACAACAGGTTGCTTCCACCGCTAGAGCTGTAGAGCCATGC	1770
Db	587	GluGlyGlnLeuGlyGluAsnAsnGluLeuLeuValGluArgLysLeuIleGluProCys	606
QY	1771	TCGCGCTAATCATCTGATATATTTTCTGTTTGGATCCGGTTATGCTTTTATTGAAACTAT	1830
Db	607	ThrAlaAsnHisLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAsnTyr	626
QY	1831	AATTTTGTAAAGATGGTAGACGCTCCGATATACAGATTGTCAGCACATTTTGTGAGCTT	1890
Db	627	AlaTyrValArgLysValProLeuAsnGluIleGluMetIleSerAlaTyrValAspLeu	646
QY	1891	AATCTAACCTCTAGAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAAGAGAG	1950
Db	647	AsnIleThrLeuLeuGluAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu	666
QY	1951	TTGCGTGTATGTTGTTGATTGATGATGAGAGTAGCTCGCGCAATCAACTACATGAA	2010
Db	667	LeuGluAspThrGlyLeuLeuAspTyrSerGluIleGlnArgAsnGlnLeuHisAla	686
QY	2011	CTTAATTTTATGACATAACAAGTAATAGAGTGATGATACAAATTTACCGGGCTGCAG	2070
Db	687	LeuLysPheTyrAspIleAspSerValValLysValAspAsnAsnValIleMetArg	706
QY	2071	GAATTCGGCTGTATGCTCTATTACTAAAAAGATGCANAAACCAATAAATGCGCCAAACCCAA	2130
Db	707	GlyIleAlaAsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValVal	726
QY	2131	TTAGAAGCAGCGCGAATGGAGTTAAACAGATCTAAATCAATGCTTAAAGCGATGACATAGCT	2190
Db	727	LeuGlyAlaAlaA-----AsnAlaValIleAlaThrValSer	738
QY	2191	TCACTACAAGACTATGCCAAG-----ATTGAA	2217
Db	739	GlyValSerSerPheLeuAsnAsnProPheGlyAlaLeuAlaValGlyLeuLeuIleLeu	758
QY	2218	GCTAGTTTATCATCTGCTTATAGTGNAGCTGAACAGTTAAACAAATTAACCTTATGCAACA	2277
Db	759	AlaGlyLeuPheAlaAlaPheLeuAlaTyrArgTyrValSerLysLeuLysSerAsnPro	778
QY	2278	TTAGAACAACTAAAAATGGCTAAACT---AATTTAGATCATGCCATCAACCAAGCTAAT	2334
Db	779	MetLysAlaLeuTyrProValThrThrLysAsnLeuLysGluSerValLysAsnGlyAsn	798
QY	2335	ACGGATAAACGACTTTTGTATTAATGAACACCCAAATTTATTGTA-----	2379
Db	799	SerGlyAsnAsnSerAspGlyGluGluAsnAspAspAsnIleAspGluLysLeuGln	818
QY	2380	-----GCATACAAAGCACTAAACACACTTTAGACAAAGCT-----	2415
Db	819	GlnAlaLysGluMetIleLysTyrMetSerLeuValSerAlaMetGluGlnGlnHis	838
QY	2416	-----GCTACTAACTTGAAGCTTTGCA	2439
Db	839	LysAlaIleLysAsnSerGlyProAlaLeuLeuAlaSerHisIleThrAsnLeuSer	858

Fri Oct 10 11:43:13 2003

QY	2440	-----TCACCTGCTTATATCAAAATTCGCAAT	2466
Db	859	LeuLysHisArgGlyProLysTyrLysArgLeuLysAsn	871
RESULT 5			
AA63230			
ID	AA63230	standard; Protein; 615 AA.	
XX	AA63230;		
AC			
XX	25-MAR-2003	(updated)	
DT	23-JUN-1995	(first entry)	
DT			
XX	Mycoplasma gallisepticum antigen (UM-67).		
DE			
XX	recombinant avipox virus; live vaccine; mycoplasma antigen.		
KW			
XX	Mycoplasma gallisepticum.		
OS			
XX	Key	Location/Qualifiers	
FT	Protein	1..615	
FT		/note= "Trp residues correspond to TGA codons"	
XX	WO9423019-A1.		
PN			
XX	13-OCT-1994.		
PD			
XX	31-MAR-1994;	94WO-JP00541.	
PF			
XX	31-MAR-1993;	93JP-0074139.	
PR			
XX	30-SEP-1993;	93JP-0245625.	
PR			
XX	(JAPG) NIPPON ZEON KK.		
PA	(SHO) SHIONOGI & CO LTD.		
PA			
XX	Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;		
PI	Saeki S, Saito, Takahashi K;		
XX	WPI; 1994-333181/41.		
DR	N-PSDB; AAQ77857.		
XX	Recombinant avipox virus combining DNA encoding a polypeptide -		
PT	exhibiting antigenicity of mycoplasma, useful for the production		
PT	of a live vaccine		
XX	Claim 4; Page 87-91; 123pp; Japanese.		
PS			
XX	A restriction fragment of the insert of M.gallisepticum genomic clone		
CC	PUM-67 containing an open reading frame was sequenced (AAQ77857). The		
CC	ORF encodes an antigenic polypeptide (AA63230). A recombinant avipox		
CC	virus comprising the coding sequence can be used as a live vaccine to		
CC	protect against infection by Mycoplasma gallisepticum.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SQ	Sequence	615 AA;	
Alignment Scores:			
Pred. No.:	4,82e-142	Length:	615
Score:	1913.00	Matches:	377
Percent Similarity:	97.46%	Conservative:	6
Best Local Similarity:	95.93%	Mismatches:	10
Query Match:	33.77%	Indels:	0
DB:	15	Gaps:	0
US-09-147-052-3 (1-3261) x AA63230 (1-615)			
QY	2050	ACAAATACCGGGCTGCAGGAATTCGGCTGTATGCTATTAATAAAGATCAAAAC	2109
Db	17	ThrSerPheLeuSerIleGlyIleSerCysMetSerIleThrLysLysAspAlaAsn	36
QY	2110	CCAAATAATGGCCAAACCAATTAGAAGCAGCGCGGAATGGAGTTAACAGATCAATCAAT	2169
Db	37	ProAsnAsnGlyGlnThrGlnLeuGlnAlaAlaArgMetGluLeuThrAspLeuIleAsn	56
QY	2170	GCTAAAGCGATGACATTAGCTTCACTACAAGACTATCCCAAGATTGAAGCTAGTTATCA	2229
Db	57	AlaLysAlaArgThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSer	76
QY	2230	TCTGCTTATAGTGAAGCTGAACAGATTAAACAATACCTTAATGCAACATAGACAACTA	2289
Db	77	SerAlaTyrSerGluAlaGluThrValAsnAsnAsnLeuAsnAlaThrLeuGluGlnLeu	96
QY	2290	AAATGGCTAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAGACT	2349
Db	97	LysMetAlaLysThrAsnLeuGluSerAlaIleAsnGlnAlaAsnThrAspLysThrThr	116
QY	2350	TTTGATAATGAACACACCCAAATTTAGTTGAACATACATAAAGCAGCTAAAACACCTTTAGAA	2409
Db	117	PheAspAsnGluHisProAsnLeuValGluAlaTyrLysAlaLeuLysThrThrLeuGlu	136
QY	2410	CAACGCTGCTACTAACCTTGAAGGTTGTCATCACTCACTTAATCAAAATCGCAATAAT	2469
Db	137	GlnArgAlaThrAsnLeuGluGlyLeuAlaSerThrAlaTyrAsnGlnIleArgAsnAsn	156
QY	2470	TTAGTGGATCTATACAATAAAGCTAGTAGTTTAATACTAAACACTAGATCCACTAAAT	2529
Db	157	LeuValAspLeuTyrAsnAsnAlaSerSerLeuIleThrLysThrLeuAspProLeuAsn	176
QY	2530	GGGGAGACGCTTTTAGATTCTTAATCAGATTACTACAGCTAATAAGATAATTAATAACG	2589
Db	177	GlyGlyMetLeuLeuAspSerAsnGluIleThrValAsnArgAsnIleAsnAsnThr	196
QY	2590	TTATCAACTATTAAATGAACAAAGACTAATGCTGATGCTATTCATCTAATAGTTTATAAA	2649
Db	197	LeuSerThrIleAsnGluGlnLysThrAsnAlaAspAlaLeuSerAsnSerPheIleLys	216
QY	2650	AAAGTGATCAAAATAAAGCAAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAA	2709
Db	217	LysValIleGlnAsnAsnGluGlnSerPheValGlyThrPheThrAsnAlaAsnValGln	236
QY	2710	CCTTCAAACTACAGTTTCTGCTTCTGCTGATGTAACACCGCTCAATTTATAAATAT	2769
Db	237	ProSerAsnTyrSerPheValAlaPheSerAlaAspValThrProValAsnTyrLysTyr	256
QY	2770	GCAAGAGGACCGTTTGGAAATGCTGATGAACCTTCAAGTAGAATTTCTTCCAAACAGAA	2829
Db	257	AlaArgArgThrValTrpAsnGlyAspGluProSerSerArgIleLeuAlaAsnThrAsn	276
QY	2830	AGTATCACAGATGTTCTTGGATTATAGTTAGTGGGAACAAACAGAAAGTACCAATTT	2889
Db	277	SerIleThrAspValSerTrpIleTyrSerLeuAlaGlyThrAsnThrLysTyrGlnPhe	296
QY	2890	AGTTTGAACACTATGCTCCATCACTGTTTATTTATTTATTTCCCTTATAGTTGGTAAA	2949
Db	297	SerPheSerAsnTyrGlyProSerThrGlyTyrLeuTyrPheProTyrLysLeuValLys	316
QY	2950	GCAGCTGATGCTAATAACGTTGGATTACATACAAAATTAATAATGGAATGTTCAACA	3009
Db	317	AlaAlaAspAlaAsnAsnValGlyLeuGlnTyrLysLeuAsnAsnGlyAsnValGlnGln	336
QY	3010	GTTGAGTTTGCCTCACTCACTAGTGCATAATAACTACAGCTAATCCACTCCAGCAGTT	3069
Db	337	ValGluPheAlaThrSerThrSerAlaAsnAsnThrThrAlaAsnProThrProAlaVal	356
QY	3070	GATGAGATTAAAGTTGCTAAATCGTTTTATCAGGTTTAAAGATTGCGCAAAACACAATC	3129
Db	357	AspGluIleLysValAlaLysIleValLeuSerGlyLeuArgPheGlyGlnAsnThrIle	376
QY	3130	GAATTAAGTGTTCACCGGTGAAGGAATATGATAAAGTTCGGCAATGATGGCAAC	3189
Db	377	GluLeuSerValProThrGlyGlyAsnMetAsnLysValAlaPrometIleGlyAsn	396
QY	3190	ATTATCTTACCTCAATGAAATAATACCTGATAGATC	3228
Db	397	IleTyrLeuSerSerAsnGluAsnAsnAlaAspLysIle	409

RESULT 6

AAR63229

ID AAR63229 standard; Protein; 610 AA.

XX AC

AAR63229;

XX DT

25-MAR-2003 (updated)

DT 23-JUN-1995 (first entry)

XX DE

Mycoplasma gallisepticum antigen (UM-66).

XX KW

recombinant avipox virus; live vaccine; mycoplasma antigen.

XX OS

Mycoplasma gallisepticum.

XX FH

Key Location/Qualifiers

FT Protein

1..610

/note= "Trp residues correspond to TGA codons"

XX PN

W09423019-Al.

XX PD

13-OCT-1994.

XX PF

31-MAR-1994; 94WO-JP00541.

XX PR

31-MAR-1993; 93JP-0074139.

XX PP

30-SEP-1993; 93JP-0245625.

XX PA

(JAPG) NIPPON ZEON KK.

XX PA

(SHIO) SHIONOGI & CO LTD.

XX PI

Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;

XX PI

Saeiki S, Saitos, Takahashi K;

XX DR

WPI; 1994-333181/41.

XX DR

N-PSDB; AAG77856.

XX PT

Recombinant avipox virus combining DNA encoding a polypeptide -

PT exhibiting antigenicity of mycoplasma, useful for the production

PT of a live vaccine

XX PS

Clam 4; Page 78-81; 123pp; Japanese.

XX CC

A restriction fragment of the insert of M.gallisepticum genomic clone

CC pUM-66 containing an open reading frame was sequenced (AAG77856). The

CC ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox

CC virus comprising the coding sequence can be used as a live vaccine to

CC protect against infection by Mycoplasma gallisepticum.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ

Sequence 610 AA;

XX

Alignment Scores:

Pred. No.: 8..6e-138 Length: 610

Score: 1859.00 Matches: 366

Percent Similarity: 96.68% Conservative: 13

Best Local Similarity: 93.37% Mismatches: 13

Query Match: 32.82% Indels: 0

DB: 15 Gaps: 0

XX

US-09-147-052-3 (1-3261) x AAR63229 (1-610)

QY

2050 ACAAAATTACGGGGCTGCAGAAATTCGGCTGTATGCTTACTTAAACAGATCAAC 2109

DB

17 ThrSerPheLeuSerIleGlyIleSerCysMetSerIleThrLysLysAspAlaAsn 36

QY

2110 CCAATATATGCCAACCCCAATTAGAACAGCGGCAATGGAGTTAACAGATCAAT 2169

DB

37 ProAsnAsnGlyInThrGlnLeuGluAlaAlaArgMetGluLeuThrAspLeuLeuAsn 56

QY

2170 GCTAAAGCGATGACATATTAGCTTCACTACAAGACTATGCCAGATTGAAGCTAGTTATCA 2229

DB

57 AlaLysAlaMetThrLeuAlaSerLeuGlnAspTyrAlaLysIleGluAlaSerLeuSer 76

RESULT 7

AAR22615

ID AAR22615

XX

AAR22615

XX

AAR22615

XX

AAR22615

XX

Db	61	ProProThrGlySerThrIleValArgLeuGluProThrArgThrCysProAspTyrHis	80
QY	211	AAAGCCACCCAGCTGGGGTGAAGAAATCCGATATTATTAAAGAGATATACAGTCCATAT	270
Db	81	LeuGlyLysAsnPhenThrGluGlyIleAlaValValTyrLysGluAsnIleAlaValTyr	100
QY	271	AAATTTAAAGTAGCGCTTTATTATAAATATCATTCAGACGACACATGACGGGACG	330
Db	101	LysPheLysAlaThrValTyrTyrLysAspValIleValSerThrAlaThrPalGlySer	120
QY	331	ACATATAGACATCACTAATCGATATACAGATAGACGCCCGTTTCCATTCAAGAGATC	390
Db	121	SerTyrThrGlnIleThrAsnArgTyrAlaAspArgValProIleProValSerGluIle	140
QY	391	ACGGATCTAATCAGCGCAAGAGAGATGCTCATCTAAAGCAAGATACCTTTAGAAACAAT	450
Db	141	ThrAspThrIleAspLysPheGlyLysCysSerSerLysAlaThrTyrValArgAsnAsn	160
QY	451	GTATATGTTCAACCGTTTGACAGGGATGCGGGNAGAAACAAGTAGTACTTCTAAACCATCA	510
Db	161	HisLysValGluAlaPheAsnGluAspLysAsnProGlnAspMetProLeuIleAlaSer	180
QY	511	AAATTTCAACACGCCCAATCTAGGGATGCGCACACGACTAATAGACAGATACCTGTGG	570
Db	181	LysTyrAsnSerValGlySerLysAlaThrHisThrAsnAspThrTyrMetValAla	200
QY	571	GGATCACCATTGATATATCGAAGCGGAACCTCCGTCATATGTATAGTAGAGCAAAATGGAT	630
Db	201	GlyThrProGlyThrTyrArgThrGlyThrSerValAsnCysIleIleGluGluValGlu	220
QY	631	GCCCGCTCTGTGTTTCGTATTCATATTTTGCATGGCCAATGGCGACATCGCAACATA	690
Db	221	AlaArgSerIlePheProTyrAspSerPheGlyLeuSerThrGlyAspIleIleTyrMet	240
QY	691	TCTCCATTTTATGTTCTATCCCAACAGAGGCTCCGCGACACCCATGGGATATCCCCAG	750
Db	241	SerProPhePheGlyLeuArg--AspGlyAlaTyrArgGluHisSerAsnTyrAlaMet	259
QY	751	GATAATTTCAACACCTAGATAGTATTTTTCATGGATTGGACAGAGGCTCGCAAGCA	810
Db	260	AspArgPheHisGlnPheGluGlyTyrArgGlnArgAspLeuAspThrArgAlaLeuLeu	279
QY	811	AGCCTTCAGTCAAGCGTAATCTTCTCATCATCACATCTTCACAGTTGGTGGGACTGG	870
Db	280	Glu---ProAlaAlaArgAsnPheLeuValThrProHisLeuThrValGlyTyrAsnTrp	298
QY	871	GCTCCAAAACACTACTCGTGTATGTTTCAATGACTAAAGTGGAAGAGGTGACTGAAATGTTG	930
Db	299	LysProLysArgThrGluValCysSerLeuValLysTrpArgGluValGluAspValVal	318
QY	931	CGTCAACAGTTAATGGGAGATACAGATTTATGCCCCGTGAACCTTCGCGCAACGTTATC	990
Db	319	ArgAspGluTyrAlaHisAsnPheArgPheThrMetLysThrLeuSerThrThrPheIle	338
QY	991	AGTAATACGACTGGATTTGATCCAAATCGCATATATTAGGACAAATGATTAAACCGGAG	1050
Db	339	SerGluThrAsnGluPheAsnLeuAsnGlnIleHisLeuSerGlnCysValLysGluGlu	358
QY	1051	GCAGAAGCAGCAATCGACAGATATTATAGGACAAATATATATGACAGTCAAGCTCAAGTT	1110
Db	359	AlaArgAlaIleAsnArgIleTyrThrArgTyrAsnSerSerHisValArgThr	378
QY	1111	GCACATGTACATATTCTTCGCTCGGGGATTTATGTAGCATATACAGCTGTGTTCTA	1170
Db	379	GlyAspIleGlnThrTyrLeuAlaArgGlyGlyPheValValPheGlnProLeuLeu	398
QY	1171	TCCAATTCCTCGCTCATATTTACCTTCAGAGAAATTTGATGAGACACAACAGGACCGATGAG	1230
Db	399	SerAsnSerLeuAlaArgLeuTyrLeuGlnGluLeuValArgGlu-----	413
QY	1231	ATGCTCGACCTGGTAAACAATPAAGCATGCAATTTATAGAAAAATGCTACCTCATCTGCA	1290
Db	414	-----AsnThrAsnHisSer---ProGlnLysHisProThrArgAsnThr	427

```

QY 1291 CGATTGCGGAGATATTCGAATGCACCAATATGAAATAAATTAACATTTAGACACACACCA 1350
Db ||| ||||| : : : |||
QY 428 ArgSerArgSerVal-----ProValGluLeuArgAlaAsnArg 441
Db ||| ||||| : : : |||
QY 1351 GCTATTAAATGACATGCTGCTCAATTCGCGATGCTCCCAATTTCTTTATCATATATA 1410
Db ||| ||||| : : : |||||
QY 442 ThrIleThrThrThrSerValGluPheAlaMetLeuGlnPheThrThrThrAspHisIle 461
Db ||| ||||| : : : |||||
QY 1411 CAACCCATATTAATGATATGTTAGTAGGATGCGACAGCTTGGTGGCAATTCGAGAT 1470
Db ||| ||||| : : : |||||
QY 462 GlnGluHisValAsnGluMetLeuAlaArgIleSerSerSerTipCysGlnLeuGlnAsn 481
Db ||| ||||| : : : |||||
QY 1471 AGAGAACTGTTTATGCGACCAAGGATAAGATTAATCCTAGGCTACACCGAGTCCA 1530
Db ||||| |||||
QY 482 ArgGluArgAlaLeuTrpSerGlyLeuPheProIleAsnProSerAlaLeuAlaSerThr 501
Db ||||| |||||
QY 1531 ACATTAGGAAGAGAGTGGCTCAAGATGTTGGGGATGCTGCTGCTATCGAGCTGC 1590
Db ||||| |||||
QY 502 IleLeuAspGlnArgValLysAlaArgIleLeuGlyAspValIleSerValSerAsnCys 521
Db ||||| |||||
QY 1591 ACTGCTATAGATCGGAATCC---GTCACCTTTGCAAAATCTATCGAGTTATCATCCTC 1647
Db ||||| |||||
QY 522 ProGluLeuGlySerAspThrArgIleIleLeuGlnAsnSerMetArgValSerGlySer 541
Db ||||| |||||
QY 1648 ACTAATACATGTTATAGCGGACCATGTTCTATTATTTTCATATGCGAAGAAACCAAGAAC 1707
Db ||||| |||||
QY 542 ThrThrArgCysTyrSerArgProLeuIleSerIleValSerLeuAsnGlySerGlyThr 561
Db ||||| |||||
QY 1708 ATACAGGCAACTCGGTGAAACCAACAGTGTCTCCACGCTAGAGCTCTAGAGCCA 1767
Db ||||| |||||
QY 562 ValGluGlnLeuGlyThrAspAsnGluLeuIleMetSerArgAspLeuLeuGluPro 581
Db ||||| |||||
QY 1768 TGCTCGGCTAATCATCGTAGATATTTCTGTTGGATCGGTTATGCTTTATTTGAAAC 1827
Db ||||| |||||
QY 582 CysValAlaAsnHisLysArgTyrPheLeuPheGlyHisHisTyrValTyrTyrGluAsp 601
Db ||||| |||||
QY 1828 TATAATTTGTTAAGATGTTAGACGCTGCGGATATACAGATTCCTAGACATTTGTCGAG 1887
Db ||||| |||||
QY 602 TyrArgTyrValArgGluIleAlaValHisAspValGlyMetIleSerThrTyrValAsp 621
Db ||||| |||||
QY 1888 CTTAATCTAACCTGCTAGAAATTTGCGTTTATCGCTTTTACACAAAGAA 1947
Db ||||| |||||
QY 622 LeuAsnLeuThrLeuLeuLysAspArgGluPheMetProLeuGlnValTyrThrArgAsp 641
Db ||||| |||||
QY 1948 GAGTTGCGTGATGTTGGTATTGATGATGAGAGTAGCTCGCGCAATCAACTACAT 2007
Db ||||| |||||
QY 642 GluLeuArgAspThrGlyLeuLeuAspTyrSerGluIleGlnArgArgAsnGlnMethHis 661
Db ||||| |||||
QY 2008 GAACCTAAATTTATGACATAACAAGTAATAGAGTGGATACAAATTTACCGGGGCTG 2067
Db ||||| |||||
QY 662 SerLeuArgPheTyrAspIleAspLysValValGlnTyrAspSerGlyThrAlaIleMet 681
Db ||||| |||||
QY 2068 CAG----- 2070
Db |||
QY 682 GlnGlyMetAlaGlnPhePheGlnGlyLeuGlyThrAlaGlyGlnAlaValGlyHisVal 701
Db |||
QY 2070 ----- 2070
Db |||
QY 702 ValLeuGlyAlaThrGlyAlaLeuLeuSerThrValHisGlyPheThrThrPheLeuSer 721
Db |||
QY 2071 ---GAATTCGGCTGATGTC----- 2088
Db ||||| |||||
QY 722 AsnProPheGlyAlaLeuAlaValGlyLeuLeuValLeuAlaGlyLeuValAlaAlaPhe 741
Db ||||| |||||
QY 2089 -----ATTCTAAAGAAGATGCAACCCA----- 2112
Db ||||| |||||
QY 742 PheAlaTyrArgTyrValLeuLysLeuLysThrSerProMetLysAlaLeuTyrProLeu 761
Db ||||| |||||
QY 2113 ---AATAATGGCCAAACCAATTAGAAGCAGCGCAATGGAGTTACAGATCTAATCAAT 2169
Db ||||| |||||
QY 762 ThrThrLysGlyLeuLysGlnLeuProGluGlyMetAspProPheAlaGluLysProAsn 781
Db ||||| |||||

```

```

QY 2170 GCTAAAGCGATGACATTACCTTCACTACAAAGACTATGCCAAGATTGAACCTAGTTTATCA 2229
Db ||| : : : |||
QY 782 AlaThrAspThrProIleGluGluIleGlyAspSerGlnAsnThrGluProSerValAsn 801
Db ||| : : : |||
QY 2230 TCTGCTTATAGT-----GAAGCTGAAACAGTTAAACAATAACCTTAAT--- 2271
Db ||| : : : |||||
QY 802 SerGlyPheAspProAspLysPheArgGluAlaGlnGluMetIleLysTyrMetThrLeu 821
Db ||| : : : |||||
QY 2272 -----GCAACATTAGAACCAATAAATGCGTAAACAACTAATTTAGAACAGCCATC 2322
Db ||| : : : |||||
QY 822 ValSerAlaAlaGluArgGlnGluSerLysAlaArgLysLysAsnLysThrSerAlaLeu 841
Db ||| : : : |||||
QY 2323 AACCAAGCTAATACGGATAAACGACTTTTGATATGAACACCAACCAATTTAGTTGAAGCA 2382
Db ||| : : : |||||
QY 841 ----- 841
QY 2383 TACAAAGCACTAAAAACCACCTTTAGAACACGCTGCTACTAACTTGAAGGTTTGTCATCA 2442
Db ||| : : : |||||
QY 842 -----LeuThrSerArgLeuThrGlyLeuAlaLeuArgAsnArg 854
QY 2443 ACTGCTTATTAATCAATTCGCAATTAATTTAGTG 2475
Db ||| : : : |||||
QY 855 ArgGlyTyrSerArgValArgThrGluIleVal 865
Db ||| : : : |||||

```

RESULT 8
AAW68404
ID AAW68404 standard; Protein; 913 AA.
AC AAW68404;
DT 28-OCT-1998 (first entry)
DE Aujeszky's disease virus glycoprotein gB.
KW Multivalent vaccine; pig; pathogen; respiratory disease; SIV; PRRSV; HCV;
KW digestive disease; Aujeszky's disease virus; pseudorabies virus; vaccine;
KW swine herpesvirus 1; swine influenza virus; hog cholera virus; vector;
KW porcine respiratory and reproductive syndrome virus; glycoprotein; SIRS;
KW swine infertility and respiratory syndrome virus;
KW Actinobacillus pleuropneumoniae.
OS Pseudorabies virus.
XX XX
XX PR2751224-AL.
XX 23-JAN-1998.
XX 19-JUL-1996; 96PR-0009338.
XX 19-JUL-1996; 96PR-0009338.
XX (INMR) RHONE MERIEUX SA.
XX WPI; 1998-112824/11.
XX N-PSDB; AAV49291.
XX Multi-valent polynucleotide vaccines against porcine pathogens -
XX consist of at least 3 plasmids able to express protective antigens
XX from specified viruses
XX Example 8; Fig 2; 63pp; French.
XX The invention relates to a multivalent vaccine for protecting pigs
XX against several pathogens, especially pathogens associated with
XX respiratory and digestive diseases. The pathogens are especially
XX selected from Aujeszky's disease virus, swine influenza virus (SIV),
XX porcine respiratory and reproductive syndrome virus (PRRSV), hog
XX cholera virus (HCV) and Actinobacillus pleuropneumoniae. The vaccines
XX are preferably composed of polynucleotide sequences encoding 3 antigens,
XX all as part of vectors. This sequence represents the Aujeszky's disease
XX virus strain NIA3 glycoprotein gB. The coding sequence was subcloned
XX into the plasmid pVR1012 to generate plasmid pAB090 for use in the
XX vaccine.

Fri Oct 10 11:43:13 2003

XX SQ Sequence 913 AA;
 Alignment Scores: 1,96e-131 Length: 913
 Pred. No.: 1779.00 Matches: 374
 Score: 61.54% Conservative: 130
 Best Local Similarity: 45.67% Mismatches: 245
 Query Match: 31.40% Indels: 70
 DB: 19 Gaps: 19

US-09-147-052-3 (1-3261) x AAW68404 (1-913)
 QY 52 GGTGACGATCCTCGAGTACCCAA-----AATGTGACATCAAGAGAGTGTTCG 105
 DB 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuaspAlaThrAlaValArg 114
 QY 106 AGCGTCAGTGTCTGAGGAAGAGTCTACGTTTATCTTTGTCCTCCACAGTGGTCA 165
 DB 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
 QY 166 ACCGTGATCGGTCTAGACCGCCGGAATGTCCCGAACCTAGAAAGCCCGAGTGG 225
 DB 133 ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGlnGlyArgAsnPhe 152
 QY 226 GGTGAGGAATCGGATATTTAAAGAGAATATCAGTCCATATAAATTTAAAGTACG 285
 DB 153 ThrGluGlyLeuAlaLeuLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
 QY 286 CTTTATTATAAAATATCATTAGACGACGACATGGAGCGGAGCATATAGACATC 345
 DB 173 IleTyrTyrLysAsnValIleValThrThrValTrpSerGlySerThrTyrAlaAlaIle 192
 QY 346 ACTAATCGATATACAGATAGACCGCCGTTTCCATTGAAGAGATCAGGATCTAATGCAC 405
 DB 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
 QY 406 GCGAAAGGAAGATGCTCTAAAGCAAGATACCTTAGAAACAATGTATGTGAAGCG 465
 DB 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnHisLysValThrAla 232
 QY 466 TTTGACAGGATGCGGGAGAAAACACAGTACTTCTAAACCATCAAAATCAACACGCC 525
 DB 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
 QY 526 GAATCTAGGCGATGCGACACACTAATGAGACGTATACCGTGTGGGATCACCATTG 585
 DB 253 GlyThrArgAlaThrPheHisThrThrAsnAspThrTyrThrLysIleGlyAlaAlaGlyPhe 272
 QY 586 TATGACAGCGGAACCTCGCTCAATGTATAGTAGAGGAATGGATGCCGCTGTGTGTT 645
 DB 273 TyrGlnThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292
 QY 646 CCGTATTCATATTTTGCATGCGCAATGGCAGATCGCGACATATCCATTTTATGTT 705
 DB 293 ProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrMetSerProPheTyrGly 312
 QY 706 CTATCCCCACAGAGCGTCCCGACAGACCCATGGGATATCCCGAGGATAATTTCAACAA 765
 DB 313 LeuArg--GluGlyAlaHisGlyGluGlnIleGlyTyrAlaProGlyArgPheGlnGln 331
 QY 766 CTAGATPAGTATTTTCAATGGATTTGGACAGCGTCCGAAAGCAAGCGCTCCAGTCAAG 825
 DB 332 ValGluHisTyrTyrProIleAspLeuAspSerArgLeuArgAlaSerGluSerValThr 351
 QY 826 CGTAACCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 885
 DB 352 ArgAsnPheLeuArgThrProHisPheThrValAlaIleThrAspIleProLysThrArg 371
 QY 886 CGTGTATGTTCAATGACTAAGTGAAGAGGTGACTGAAATGTCGGT---GCAACAGTT 942
 DB 372 ArgValCysSerLeuAlaLysTrpArgGluAlaGluGluMetThrArgAspGluThrArg 391

QY 943 AATGGGATACAGATTTATGGCCGCGTAACTTTCGGCAAGCTTTTATCAGTAATACGACT 1002
 DB 392 AspGlySerPheArgPheThrSerArgAlaLeuGlyAlaSerPheValSerAspValThr 411
 QY 1003 GAGTTTGATCAAAATCGCATCATATTAGGCAATATGTTAAACCGGAGGAGCAAGCAGCA 1062
 DB 412 GlnLeuAspLeuGlnArgValHisLeuGlyAspCysValLeuArgGluAlaSerGluAla 431
 QY 1063 ATCGAGCAGATATTTAGGACAAAATATATGACAGTCAAGTCAAGGTTGGGA---CATGTA 1119
 DB 432 IleAspAlaIleTyrArgArgTyrAsnSerThrHisValLeuAlaGlyAspArgPro 451
 QY 1120 CAATATTTCTTGGCTCTCGGGGATTTATGTAGCATATCAGCCGTGTTCTATCCAAATCC 1179
 DB 452 GluValTyrLeuAlaArgGlyGlyPheValValAlaPheArgProLeuIleSerAsnGlu 471
 QY 1180 CTGGCTCATATCTACCTCAGAGAAATGATGAGAGACAACAGACCGATGAGATGTCGAC 1239
 DB 472 LeuAlaGlnLeuTyrAlaArgGluLeu-----GluArgLeuGly 484
 QY 1240 CTGGTAAACAATAAGCATGCAATTTATAAGAAAAATGCTACCTATTGTACGATTCGGG 1299
 DB 485 LeuAlaGly-----ValValGlyProAlaAlaProAlaAlaAlaArgAlaArg 501
 QY 1300 CGAGAT-----ATTCGAAATGCCCAATAGA 1326
 DB 502 ArgSerProGlyProAlaGlyThrProGluProProAlaValAsnGlyThrGlyHisLeu 521
 QY 1327 AAAATAACATTAGACACACACAGCATCTAAATCGCATCTCTGTTCAATTCGCCATG 1386
 DB 522 ArgIleThr-----ThrGlySerAlaGluPheAlaArg 532
 QY 1387 CTCCAATTTCTTATGATCATATACAAACCATATTAATGATATGTTTAGTGGATTGCC 1446
 DB 533 LeuGlnPheThrTyrAspHisIleGlnAlaHisValAsnAspMetLeuGlyArgIleAla 552
 QY 1447 ACAGCTGTGCGCAATCGAGATAGAACITGTTTATGTCACCAAGGATAAAGAT 1506
 DB 553 AlaAlaTrpCysGluLeuGlnAsnLysAspArgThrLeuTrpSerGluMetSerArgLeu 572
 QY 1507 AATCTAGCGCTACACGAGTGCACATTTAGGAAGAGAGATGGCTGCAAGATGTTGGGG 1566
 DB 573 AsnProSerAlaValAlaThrAlaAlaLeuGlyGlnArgValCysAlaArgMetLeuGly 592
 QY 1567 GATGTCGCTCTGTATCGGCTCCACCTGATATAGATCCGGAATCCGTCACATTTGCAAAAT 1626
 DB 593 AspValMetAlaIleSerArgCysValGluValArgGly---GlyValTyrValGlnAsn 611
 QY 1627 TCTATGCGATTTATACATCCACTATACATGTTATAGCGGACCATTTGTTCTATTTC 1686
 DB 612 SerMetArgValProGlyGluArgGlyThrCysTyrSerArgProLeuValThrPheGlu 631
 QY 1687 TATGGAGAAACCAAGGAACATACAGGACACTCGGTGCAAAACACAGAGTGTGCTTCCA 1746
 DB 632 His---AsnGlyThrGlyValIleGluGlyGlnLeuGlyAspAsnGluLeuLeu 650
 QY 1747 ACGCTAGAGCTGTAGAGCCATCTCGGCTAATCATCTAGATATATTTCTGTTGGATCC 1806
 DB 651 SerArgAspLeuIleGluProCysThrGlyAsnHisArgTyrPheLysLeuGlySer 670
 QY 1807 GGTATGCTTTTGAATACTATAATTTTGAAGATGGTAGAGCGTCCGATATACAG 1866
 DB 671 GlyTyrValTyrTyrGluAspTyrAsnTyrValArgMetValGluValPro-----Glu 688
 QY 1867 ATTGCTAGCATTGTCGAGCTTAATCTAACCCCTGCTAGAGATCCGGAATTTGCGCT 1926
 DB 689 ThrIleSerThrArgValThrLeuAsnLeuThrLeuLeuGluAspArgGluPheLeuPro 708
 QY 1927 TTATCCGTTTACAAAAGAGAGTGTGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1986
 DB 709 LeuGluValTyrThrArgGluGluLeuAlaAspThrGlyLeuLeuAspTyrSerGluIle 728
 QY 1987 GCTCGCCGCACTCACTACATGAACTTAAATTTTATGACATAAACAANGTANTAGAGTG 2046

```

Db 729 GlnArgArgAsnGlnLeuHisAlaLeuLysPheTyrAspIleAspArgValValIysVal 748
QY 2047 GATCAAAAT-----TACGGCGGCTCCAGGAATTC 2076
Db 749 AspHisAsnValValLeuArgGlyIleAlaAsnPhePheGlnGlyLeuGlyAspVal 768
QY 2077 GGCTGT-----ATGCTCTATTACTAAAGATGCAACCCAAATAT 2118
Db 769 GlyAlaAlaValGlyLysValValLeuGlyAlaThrGlyAlaValIleSerAlaValGly 788
QY 2119 GCCAAACCCATTAAGACGCGCGAATGAGTTAACAGATCTAATCAATGCTAAAGCG 2178
Db 789 GlyMetValSerPhe-----LeuSerAsnPropPheGlyAlaLeuAla 802
QY 2179 ATGACATTAGTTCACCTACAGACTATGCCAAGATTGAAGTAGTTATCATCTGCTTAT 2238
Db 803 IleGlyLeuLeuValLeu-----AlaGlyLeuValAlaAlaPhe---LeuAlaTyr 818
QY 2239 AGTGAAGCTGAACAGTTAAACATATACCTTAAATGCAACATTAGAACAACTAAATGGCT 2298
Db 819 ArgHisIleSerArgLeuArgArgAsnPrometIysAlaLeuTyrProVal---ThrThr 837
QY 2299 AAACTAATTTAGATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAAT 2358
Db 838 LysThrLeuLysGluAspGlyValAspGluGlyAspValAspGluAlaLysLeuAspGln 857
QY 2359 GAACACCAAAATTTAGTTGAAGCATACAAACGACATAAAACACCTTTTGAACAACGT 2415
Db 858 AlaArgAspMetIle-----ArgTyrMetSerIleValSerAlaLeuGluGlnGln 874

RESULT 9
AAP50035
ID AAP50035 standard; Protein; 845 AA.
AC AAP50035;
XX
XX
XX 25-MAR-2003 (updated)
DT 07-SEP-1991 (first entry)
XX
XX N-terminal sequence of the pseudorabies virus (PRV) gII protein.
XX Sub-unit vaccine; antigen; immunoreactive determinant.
XX
XX Pseudorabies virus (PRV).
OS
FH Key Location/Qualifiers
FT Modified-site 151..153 /label- potential glycosylation site
FT Modified-site 261..263 /label- potential glycosylation site
FT Modified-site 441..443 /label- potential glycosylation site
FT Modified-site 516..518 /label- potential glycosylation site
FT Modified-site 573..575 /label- potential glycosylation site
FT Modified-site 633..635 /label- potential glycosylation site
FT Modified-site 697..699 /label- potential glycosylation site
XX
XX EP162738-A.
PN
XX
XX 27-NOV-1985.
PD
XX
XX 09-APR-1985; 85EP-0400704.
PF
XX
XX 03-APR-1985; 85US-0719773.
PR
XX 09-APR-1984; 84US-0598073.
PA
XX (MOLE-) MOLECULAR GENETICS INC.
XX

```

```

PI Robbins AK, Watson RJ, Enquist LW;
DR WPI; 1985-298083/48.
DR N-PSDB; AAN50036.
XX
XX Prodn. of pseudorabies virus sub-unit vaccines - useful for
PT conferring protection against the virus infections and for
PT diagnosis esp. with sheep sera
XX
XX Example; Fig 13; 153pp; English.
XX
XX The inventors claim a recombinant vector comprising a DNA sequence
CC coding for an immunoreactive and antigenic determinant of a
CC pseudorabies virus protein, and a polypeptide having an
CC immunoreactive and antigenic determinant of a pseudorabies virus
CC protein. Vaccines contg. the polypeptide are obtd. economically and
CC in large amts., for use in conferring protection against
CC pseudorabies virus.
CC (Updated on 25-MAR-2003 to correct PD field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 845 AA;

Alignment Scores:
Pred. No.: 2,21e-130 Length: 845
Score: 1765.50 Matches: 351
Percent Similarity: 65.91% Conservative: 111
Best Local Similarity: 50.07% Mismatches: 188
Query Match: 31.17% Indels: 51
DB: 6 Gaps: 13

US-09-147-052-3 (1-3261) x AAP50035 (1-845)
QY 52 GGTACGAACCTCATCTCCGAGTACCCAA-----AATGTGACATCAAGAGAACTGTTTCG 105
Db 95 GlyProSerGluAlaProAspGlyGluTyrGlyAspLeuAspAlaArgThrAlaValArg 114
QY 106 AGCGTCCAGTTGCTGAGAGAGAGTCTACGTTTTTATCTTTTGTCCCGCCAGTGGTTCA 165
Db 115 AlaAla-----AlaThrGluArgAspArgPheTyrValCysProProSerGlySer 132
QY 166 ACCGTGATCGCTGTAGAACCGCGCGAAAGTCCCGAACCTAGAAAGCCACCAGTGG 225
Db 133 ThrValValArgLeuGluProGluGlnAlaCysProGluTyrSerGluTyrArgAsnPhe 152
QY 226 GGTGAAGGAATCGCGATATTATTTAAAGAGATATCAGTCCATATAAATTTAAAGTACG 285
Db 153 ThrGluGlyIleAlaValLeuPheLysGluAsnIleAlaProHisLysPheLysAlaHis 172
QY 286 CTTTATTATAAATATATTCATTCAGACGAGCATGGACGGGACGACATATAGACAGATC 345
Db 173 IleTyrTyrLysAsnValIleValThrThrValTyrPserGlySerThrTyrAlaAlaIle 192
QY 346 ACTAATCATGATACAGATAGGACGCGCTTTCATTCATTCAGAGATCAGGATCTAATCGAC 405
Db 193 ThrAsnArgPheThrAspArgValProValProValGlnGluIleThrAspValIleAsp 212
QY 406 GCGAAAGGAGATGCTCATCTAAACCAAGATACCTTAGAACAAATGTATATGTTGAAGCG 465
Db 213 ArgArgGlyLysCysValSerLysAlaGluTyrValArgAsnHisLysValThrAla 232
QY 466 TTGACAGGGATGCGGGAGAGAAAACAGTACTTCTAAACCATCAAAATTCACAGCCGC 525
Db 233 PheAspArgAspGluAsnProValGluValAspLeuArgProSerArgLeuAsnAlaLeu 252
QY 526 GAATCTAGGCGATGGCACACGACTAATGAGAGGTATACCGTGTGGGGATCCACCTGGATA 585
Db 253 GlyThrArgGlyTyrPheHisThrThrAsnAspThrThrThrLysIleGlyAlaGlyPhe 272
QY 586 TATCGAAGCGGAACCTCGCTCAATTTGTATAGTAGAGAAATGAGTCCCGCTCTCTGTTT 645
Db 273 TyrHisThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyr 292

```

646 CGGTATTCATATTTGCAATGCGCAATGCGGACATCGGACACATATCTCCATTTTATGGT 705
 293 ProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyrMetSerPheProGly 312
 706 CTATCCCGACGAGCTCGGACAGCCATCGGATATCCAGGATAATTTCAACAA 765
 313 LeuArg--GluGlyAlaHisGlyGluHisIleGlyTyrAlaProGlyArgPheGlnGln 331
 766 CTAGATAGCTATTTTCAATGATTTGGACAGCGTCGGAAGCGTCGAGCTCCAGTCAG 825
 332 ValGluHisTyrTyrProIleAspLeuAspSerArgLeuArgAlaSerGluSerValThr 351
 826 CGTAATCTTCATCATACATCAGCTTCACAGTGGGTGGGACTGGGCTCCAAACTACT 885
 352 ArgAsnPheLeuArgThrProHisPheThrValAlaIlePhePhePhePhePhePhe 371
 886 CGGTATGTTCAATGACTAAGTGAAGAGGCTGACTGAATGTTGCGT---GCAACAGTT 942
 372 ArgValCysSerLeuAlaLysTyrPargGluAlaGluMetThrArgAspGluThrArg 391
 943 AATGGGAGATACAGATTTATGGCCGCTGAACTTCGCGCAACGTTTATCATGTAATCAGCT 1002
 392 AspGlySerPheArgPheThrSerArgAlaLeuGlyAlaSerPheValSerAspValThr 411
 1003 GAGTTTTCATCAATCGCATCATATAGGACAATGATTAACGCGAGGCGAGCAGCA 1062
 412 GlnLeuAspLeuGlnArgValHisLeuGlyAspCysValLeuArgGluAlaSerGluAla 431
 1063 ATCGAGCAGATATTTAGGACAAATATATATGACAGCTCAGCTCAAGTTGGA---CATGTA 1119
 432 IleAspAlaIleTyrArgArgTyrAsnSerThrHisValLeuAlaGlyAspArgPro 451
 1120 CAATATTTCTGGCTCGGGGATTTATGTAGCATATCAGCTGTTCTATCCAAATCC 1179
 452 GluValTyrLeuAlaLysTyrGlyPheValValAlaPheArgProLeuIleSerAsnGlu 471
 1180 CTGGCTCATATGTACCTCAGAGATTTGATGAGAGCAACAGGCGGATGATGCTGCAC 1239
 472 LeuAlaGlnLeuTyrAlaArgGluLeu-----GluArgLeuGly 484
 1240 CTGGTAACATACCATGCAATTTATTAAGAAAAATGCTACCTGTTGTCACCATTCGG 1299
 485 LeuAlaGly-----ValValCysProAlaAlaProAlaAlaAlaArgAlaArg 501
 1300 CGAGAT-----ATTCGAATGACCAACATAGA 1326
 502 ArgSerProGlyProAlaGlyThrProGluProAlaValAsnGlyThrGlyHisLeu 521
 1327 AAAATAACATTAGACGACACACCATGATTAATCGACATGCTGTTCAATTCGCCATG 1386
 522 ArgIleThr-----ThrGlySerAlaGluPheAlaArg 532
 1387 CTCCAATTTCTTTATGATCATATACAAACCCATATTAATGATGTTTATGAGTATGCC 1446
 533 LeuGlnPheThrTyrAspHisIleGlnAlaHisValAsnAspMetLeuGlyArgIleAla 552
 1447 ACAGCTTGTCGGAATTCAGAAATAGAGAACTGTTTATGCGCAGGAGGATTAAGATT 1506
 553 AlaAlaTyrCysGluLeuGlnAsnLysAspArgThrLeuTyrSerGluMetSerArgLeu 572
 1507 AATCTTAGCGCTACAGGAGTCAACATTAGGAGGAGTGGCTGCAAGAGATGTTGGGG 1566
 573 AsnProSerAlaValAlaAlaLeuAlaLeuGlyGlnArgValSerAlaArgMetLeuGly 592
 1567 GATGTCGCTGCTATCGAGCTGCACATGCTATAGATGCGGAATCCGTCACCTTGCATAAT 1626
 593 AspValMetAlaIleSerArgCysValGluValArgGly---GlyValTyrValGlnAsn 611
 1627 TCATGCGAGTATATACATCCACTATATGATGTTATAGCCGACCATGTTGTTCTATTTC 1686
 612 SerMetArgValProGlyGluArgGlyThrCysTyrSerArgProLeuValThrPheGlu 631
 1687 TATGGAGAAACCAAGAAACATACAGGAGCAACTCGGTGAAACCAACAGTGTCTTCCA 1746

632 His---AsnGlyThrGlyValIleGluGlyGlnLeuGlyAspAsnGlnLeuLeuIle 650
 1747 ACCTAGAGCTCTAGAGCCATCTCGGCTAATCATCTAGATATTTCTGTTTGGATCC 1806
 651 SerArgAspLeuIleGluProCysThrGlyAsnHisArgTyrPheLysLeuGlySer 670
 1807 GGTATGCTTTATTTGAAACTATATAATTTTGAAGATGGTAGCGCTGCCGATATACAG 1866
 671 GlyTyrValTyrTyrGluAspTyrAsnTyrValArgMetValGluValPro-----Glu 688
 1867 ATTCTAGACACATTTGCGAGCTTAATCTAACCTGCTAGACATCGGAAATTTTGCCT 1926
 689 ThrIleSerThrArgValThrLeuAsnLeuThrLeuLeuGluAspArgGluPheLeuPro 708
 1927 TTATCCGTTTACAAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1986
 709 LeuGluValTyrThrArgGluGluLeuAlaAspThrGlyLeuLeuAspTyrSerGluIle 728
 1987 GCTCGCCCAATCACTACATGAACTTAATTTTATGACATATAACAAGTAATAGAGTG 2046
 729 GlnArgArgAsnGlnLeuHisAlaLeuLysPheTyrAspIleAspArgValValLysVal 748
 2047 GATACAAAT-----TACGCGGGCTGCAGGAATTC 2076
 749 AspHisAsnValValLeuLeuArgGlyIleAlaAsnPhePheGlnGlyLeuGlyAspVal 768
 2077 GGC 2079
 769 Gly 769
 RESULT 10
 AAR27807
 ID AAR27807 standard; Protein; 933 AA.
 XX AC AAR27807;
 XX DT 10-MAR-1993 (first entry)
 XX DE Bovine herpes virus type 1 glycoprotein I.
 XX KW BHV-1; vaccine; gi.
 XX OS Bovine herpes virus type 1.
 XX FH Key Location/Qualifiers
 XX FT Peptide 1..505 /note="gIb peptide"
 XX FT Peptide 506..933 /note="gIc peptide"
 XX FT Region 767..828 /note="putative transmembrane sequence"
 XX US5151267-A.
 XX PN 29-SEP-1992.
 XX PF 15-JUL-1988; 88US-0219939.
 XX PR 15-JUL-1988; 88US-0219939.
 XX PA (UYSA-) UNIV SASKATCHEWAN.
 XX PI Babiuk L, Fitzpatrick D, van den Hurk S, Zamb T;
 XX DR WPI; 1992-348566/42.
 XX DR N-PSDB; AAQ29515.
 XX PT Protecting bovine hosts from bovine herpes virus type 1 infection -
 XX PT using vaccine contg. polypeptide neutralising epitope of GI, GII or
 XX PT GIII BHV-1 glyco:protein
 XX PS Disclosure; Fig 5; 48pp; English.

xx The sequence is that of the BHV-1 glycoprotein gI which comprises the
cc gIa/gIb/gIc complex, the glycoprotein maintains the proper epitopes
cc necessary to fully protect immunised animals from disease. It elicits
cc antibodies that neutralise virus infectivity and induce complement-
cc mediated cell lysis. It can be used in vaccine compns. which are
cc more protective than killed virus and attenuated live-virus vaccines.
cc See also AAR27808 and AAR27809.

AA	Sequence	933 AA;
SQ		

Alignment Scores:		
Pred. No.:	7,84e-129	933
Score:	1746.00	341
Percent Similarity:	49.27%	Conservative: 106
Best Local Similarity:	45.79%	Mismatches: 218
Query Match:	30.82%	Indels: 20
DB:	13	Gaps: 6

US-09-147-052-3 (1-3261) x AAR27807 (1-933)

[illegible]

Qy	832	TTTCTCATCACYCACATTCACAGATTGGGTGGGACGCGGCTCCAAAAACTACTCTGGTA	891
Db	356	PheLeuArgThrGlnHisValThrValAlaTrpAspTrpValProLysArgLysAsnVal	375
Qy	892	TGTTCAATGACTAGTGGAAGAGGTGCATGAAATGTTGCGTGCACAGATTAAATGGAGA	951
Db	376	CysSerLeuAlaLysTrpArggLuLaAspLumetLeuArgaspGlusSerArggLyasn	395
Qy	952	TACAGATTATGGCCCCGTGAACCTTCGCCAACGTTTTATCAGTAATCAGCACTAGAGTTGAT	1011
Db	396	PheArgPheThrAlaArgSerLeuSerAlaThrPheValSerAspSerHisThrPheAla	415
Qy	1012	CCAANTCGCATCATATTAGGACAATGTATTAAACGCGAGGAGGAGCAAGCAATCGAGCAG	1071
Db	416	LeuGlnAsnValProLeuSerAspCysValilegluLaIlegluLaAlaValgluArg	435
Qy	1072	ATATTTAGGACAAAATAATATGACAGTCACGTCACAGGTTGGACATGTACAAATATTCTTG	1131
Db	436	ValTyArgGluArgTyAsnGlyThrHisValLeuSerGlySerLeuGluThrTyLeu	455
Qy	1132	GCTCTCGGGGATTTATGTAGCATATACGCTTCTCTCCAAATCCCTGGCTCATATG	1191
Db	456	AlaArgGlyGlyPheValValAlaPheArgPrometLeuSerAsnGluLeuAlaLysLeu	475
Qy	1192	TACTCAGACAATGTATGAGACACACAGGACGATGAGATGCTGCGACCTGGTAAACAAT	1251
Db	476	TyrLeuGlnGluLeuAlaArgSerAsnGlyThrLeuGluGlyLeuPheAlaAla-----	493
Qy	1252	AAGCATGCAATTTATAGAAAATGCTTACCTCATTTGTCAGNTTGGCGGAGATATTCGA	1311
Db	494	-----AlaAlaProLysProGlyPro-----ArgArgAlaArgArgAlaAlaPro	508
Qy	1312	AATGCACCAAATAGAAAAATAACATTA-----GAGCAGACCACA	1350
Db	509	SerAlaProGlyGlyProGlyAlaAlaAsnGlyProAlaGlyAspGlyAspAlaGlyGly	528
Qy	1351	GCTATTAAATCGACATCGTCTGTTCATTCGCCATGCTCCAAATCTTTTATGATCAPATA	1410
Db	529	ArgValThrValSerAlaGluPheAlaAlaLeuGlnPheThrTyRasPhisIle	548
Qy	1411	CAAAACCATATTAATGATATGTTTAGTAGATGCCACAGCTGGTCGGAATTCACAAT	1470
Db	549	GlnAspHisValAsnThrMetPheSerArgLeuAlaThrSerTrpCysLeuLeuGlnAsn	568
Qy	1471	AGAGAACCTGTTTTATGCGCAGGAGGATAAAGATTAATCCTAGCGCTACAGCGAGTCA	1530
Db	569	LysGluArgAlaLeuTrpAlaGluAlaLysLeuAsnProSerAlaAlaAlaSerAla	588
Qy	1531	ACATTAGGAGGAGATGGCTGCAACAGATGTGGGGATGCGCTGCTGATCGAGCTGC	1590
Db	589	AlaLeuAspArgArgAlaAlaAlaArgMetLeuGlyAspAlaMetAlaValThrTyR-Cys	608
Qy	1591	ACTGCTATAGATCGGAATCGCTCACCTTCMAAATTCATGCGAGTTATCATCCACT	1650
Db	609	HisGluLeuGlyGluGlyArgValPheIleGlnAsnSerMetArg-----AlaProGly	626
Qy	1651	AATACATGTTATAGCCGACCATTCGGTCTATTTTTCATATGAGAAAAACCAAGGAACATA	1710
Db	627	GlyValCysTySerArgProProValSerPheAlaPheGlyAsnGluSerGluProVal	646
Qy	1711	CAGGCAACTCGGTGAAAAACACAGCTTGCTTCCACGCTAGAGGCTAGAGCCATGC	1770
Db	647	GlucGlnLeuGlyGluAspAsnGluLeuLeuProGlyArgGluLeuValGluProCys	666
Qy	1771	TCGCTTAATCATCTAGATATTTCTGTTGGATCCGGTTATGCTTATTATTGAAACTAT	1830
Db	667	ThrAlaAsnHisLysArgTyPheArgPheGlyAlaAspTyValTyTyGluAsnTyR	686
Qy	1831	AATTTTGTAGATGGTAGCGCTGCGCSNATACAGATTGCTAGCACATTTGCGAGCTT	1890
Db	687	AlaTyValArgArgValProLeuAlaGluLeuGluValIleserThrPheValAspLeu	706
Qy	1891	AATCTAAACCTGCTAGAAATCGCGAAATTTTGCTTTATCCGTTTACACAAAAGAGAG	1950

Db 707 AsnLeuThrValLeuGlnAspArgGluPheLeuProLeuGluValTyrThrArgAlaGlu 726
 QY 1951 TTGCGTGATGTTGGTGTATTGATTATGCAGAGTGTGCTCCCGCAATCAACTACATGAA 2010
 Db 727 LeuAlaAspThrGlyLeuLeuAspTyrSerGluLeuGlnArgArgAsnGlnLeuHisGlu 746
 QY 2011 CTTAAATTTTATGACATAAACAAGTAATAGAGTGGATACAAATTACGCG----- 2061
 Db 747 LeuArgPheTyrAspIleAspArgValValTyrThrAspGlyAsnMetAlaIleMetArg 766
 QY 2062 GGGCTGCAGGAATTC 2076
 Db 767 GlyLeuAlaAsnPhe 771

RESULT 11
 AAR41343
 ID AAR41343 standard; Protein: 933 AA.
 XX AAR41343;
 AC AAR41343;
 DT 18-FEB-1994 (first entry)
 XX Bovine herpesvirus type I g1 glycoprotein.
 DE BHV-1; Bovine herpesvirus; vaccine; neutralising epitope;
 KW glycoprotein; coat protein; g1; herpes; virus.
 KW Bovine herpesvirus.
 OS

XX Key Location/Qualifiers
 FH Cleavage-site 505
 FT /note= g1a precursor protein cleaved to give g1b
 FT and g1c
 FT Region 767..828
 FT /label= Putative transmembrane region.
 XX CA2057387-A.
 PN 12-JUN-1993.
 XX 11-DEC-1991; 91CA-2057387.
 PF 11-DEC-1991; 91CA-2057387.
 PR 11-DEC-1991; 91CA-2057387.
 XX (VETE-) VETERINARY INFECTIOUS DISEASE.

PA
 XX Babluk L, Fitzpatrick D, Van DER HURK S, Zamb T;
 PI WPI; 1993-273267/35.
 DR P-PSDB; AAR41343.
 XX Recombinant bovine herpes type 1 proteins g1, gIII and gIV - and
 PT DNA encoding them, for use in vaccines against e.g. shipping
 PT fever
 XX Claim 2; Figure 5; 154pp; English.
 PS

XX The bovine herpesvirus g1 glycoprotein (BHV-1) is used in the
 CC construction of vectors which are subsequently used to transform
 CC host cells. g1 produced by these cells can then be used to produce
 CC subunit vaccines comprising one or more neutralising epitopes of the
 CC g1 glycoprotein. The subunit vaccines are used to protect cattle
 CC from disease, especially shipping fever which is a complex syndrome
 CC which often includes infection by BHV-1. They are substantially
 CC more protective than previous killed or live attenuated virus
 CC vaccines.
 XX SQ Sequence 933 AA;

Alignment Scores: 7,84e-129 Length: 933
 Pred. No.: 1746.00 Matches: 341
 Score:

Percent Similarity: 65.26% Conservative: 106
 Best Local Similarity: 49.78% Mismatches: 218
 Query Match: 30.82% Indels: 20
 DB: 14 Gaps: 6
 US-09-147-052-3 (1-3261) x AAR41343 (1-933)
 QY 52 GGTACGAACTCATCTCCGAGTACCCAAATGTGACATCAAGAGAAAGTTGTTTCGAGCGTC 111
 Db 97 GlyAspAspAlaAlaSerProAspAsnSerThrAspValArgAlaAlaLeuArgLeuAla 116
 QY 112 CAGTGTCTGAGGAGAGTCTAGCTTTTATCTTGTCCCCACAGTGGTTCACACCGTG 171
 Db 117 GlnAlaAlaGlyGluAsnSerArgPhePheValCysProProSerGlyAlaThrVal 136
 QY 172 ATCGTCTAGAACCGCGCGAAATGTCCGAACTAGAAAAACCCACCGAGGGTGAA 231
 Db 137 ValArgLeuAlaProAlaArgProCysProGluTyrGlyLeuGlyArgAsnTyrThrGlu 156
 QY 232 GGAATCGGGATATTATTAAAGAGAAATATCAGTCCATATAAATTTAAAGTGACGCTTAT 291
 Db 157 GlyIleGlyValIleTyrLysGluAsnIleAlaProTyrThrPheLysAlaTyrIleTyr 176
 QY 292 TATAAAATATCATTCAGACGACGACATGGACGGGACGACATATAGACATATAGACATCAAT 351
 Db 177 TyrLysAsnValIleValThrThrTrpAlaGlySerThrTyrAlaAlaIleThrAsn 196
 QY 352 CGATATACAGATAGACGCCGTTTCCATTGAAAGAGATCAGGATCTAATCGACGGCAAA 411
 Db 197 GlnTyrThrAspArgValProValGlyMetGlyGluIleThrAspLeuValAspLysLys 216
 QY 412 GGAAGATCTCATCTAAAGCAAGATCCTTAGAACCAATGTATATGTTGAAGCGTTGAC 471
 Db 217 TrpArgCysLeuSerLysAlaGluTyrLeuArgSerGlyArgLysValValAlaPheAsp 236
 QY 472 AGGGATGCGGAGAAAAACAAGTACTTCTAAACCATCAAAATTCACACGCCCGAATCT 531
 Db 237 ArgAspAspProTyrGluAlaProLeuLysProAlaArgLeuSerAlaProGlyVal 256
 QY 532 AGGGATGCGCACACGACTAATGAGACGTATACCGTGTGGGATCACCATGATATATCGA 591
 Db 257 ArgGlyTrpHisThrAspAspValTyrThrAlaLeuGlySerAlaGlyLeuTyrArg 276
 QY 592 ACGGGAACCTCGCTCAATGTATAGTAGAAATGGATGCCGCTGTGTGTTCCGPTAT 651
 Db 277 ThrGlyThrSerValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 296
 QY 652 TCATATTTTGAATGGCCAAATCGGACATCGGACATATCTCCATTTTATGGTCTATCC 711
 Db 297 AspSerPheAlaLeuSerThrGlyAspIleIleTyrMetSerProPheTyrGlyLeuArg 316
 QY 712 CCACAGAGGCTCCGCGACAGACCCATGGGATATCCCGAGTAATTTCAAACAACATAGAT 771
 Db 317 ---GluGlyAlaHisArgGluHisThrSerTyrSerProGluArgPheGlnIleGlu 335
 QY 772 AGCTATTTTCAATGGATTGGCAAGCGCTCGAAAGAACGCTTCCAGTCAAGCGTAAC 831
 Db 336 GlyTyrTyrLysArgAspMetAlaThrGlyArgLeuLysGluProValSerArgAsn 355
 QY 832 TTCTCATCATCACATCACATTCACAGTGGTGGGCTGGCTCCCAAAACTACTCGTGA 891
 Db 356 PheLeuArgThrGlnHisValThrValAlaIleTrpAspIleValProLysArgLysAsnVal 375
 QY 892 TGTTCATGACTAAGTGGAAAGAGGTGACTGAAATGTTGCGTGCAACAGCTTAATGGAGA 951
 Db 376 CysSerLeuAlaLysTyrPargGluAlaAspGluMetLeuArgAspLysSerArgGlyAsn 395
 QY 952 TACAGATTTATGGCCGTGAACCTTTCGGCAACCTTTATCAGTAATACGACTGAGTTGAT 1011
 Db 396 PheArgPheThrAlaArgSerLeuSerAlaThrPheValSerAspSerHisThrPheAla 415
 QY 1012 CCAAATCGCATCATATTAGGACAATGTATTAAACGGGAGGACAGCAAGCAATCGAGCAG 1071

Fri Oct 10 11:43:13 2003

CC stable and more immunogenic than known modified live vaccines,
CC providing better protection against development of clinical signs
CC and viral shedding.

SQ Sequence 943 AA;

Alignment Scores:

Pred. No.: 1.35e-128 Length: 943
Score: 1743.00 Matches: 377
Percent Similarity: 57.54% Conservative: 161
Best Local Similarity: 40.32% Mismatches: 267
Query Match: 30.77% Indels: 130
DB: 21 Gaps: 21

US-09-147-052-3 (1-3261) x AAY32470 (1-943)

QY 7 TATTTAGCGGAATTCGATATT----- 30
DB 22 TyrPheArgGlnAArgCysPheProSerLeuLeuGlyLeuAlaAlaThrGlySerArg 41
QY 31 -----TTCTTATAGTT 42
DB 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
QY 43 ATTCTATAT-----GGTAGCACTCATCTCCG 69
DB 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
QY 70 ACTACCCAAATGTGACATCAAGAGATT----- 99
DB 82 ProArgTyrThrValAlaThrProGluValGlyValHisGlnAsnGlnLeuGlnIle 101
QY 100 -----GTTTCGAGCGTCCAGTGTCTGAGGAAGAG----- 129
DB 102 ProProIleCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
QY 130 TCTAGCTTTATCTTTGTCCTCCACCAGTGGTTCACCGTCAACCGTCAACCGTCAACCGCG 189
DB 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
QY 190 CGAAATGTCGCAACCTAGAAAGCCAGCGAGTGGGTGAAGGAATCGGATATTATT 249
DB 142 ArgAlaCysProAspTyrLysLeuGlyLysAsnPheThrGluGlyIleAlaValIlePhe 161
QY 250 AAAGAGATATCAGTCCATATAATTAAGTACGCGCTTTATATAAAATATCATCAG 309
DB 162 LysGluAsnIleAlaProTyrLysPheLysAlaAsnIleTyrTyrLysAsnIleIleMet 181
QY 310 ACAGCAGATGCGGGGAGCAGATATACAGATCACTAATCGATACAGATAGGAGC 369
DB 182 ThrThrValTrpSerGlySerTyrAlaValThrThrAsnArgTyrThrAspArgVal 201
QY 370 CCGTTCATTAAGAGATACGCGATCTAATCGAGCGCAAGAGAGATGCTCATATAA 429
DB 202 ProValLysValGlnGluIleThrAspLeuIleAspArgGlyMetCysLeuSerLys 221
QY 430 GCAAGATACCTTAGAACCAATGTATATGTGAAGCGTTTACAGAGGATCGGGAGAAA 489
DB 222 AlaAspTyrValArgAsnAsnTyrGlnPheThrAlaPheAspArgAspGluAspProArg 241
QY 490 CAAGTACTTCTAAACCATCAAAATCAACACCGCCGAATCTAGGCGATGCGACAGACT 549
DB 242 GluLeuProLeuLysProProSerSerThrLeuSerArgValArgGlyTrpHis---Thr 260
QY 550 AATGAGACGTATACCGTGTGGGATCACCATTGATATATCGAACGGGAACCTCCGTCAT 609
DB 261 AsnGluThrTyrThrLysIleValLeuLeuAspPheHisSerGlyThrSerValAsn 280
QY 610 TGATATAGAGAAATGGATGCGCGTCTGTGTTCCGTATTCATATTTGCAATGGCG 669
DB 281 CysIleValGluValAspAlaArgSerValTyrProTyrAspSerPheAlaIleSer 300
QY 670 AATGCGCAGATCGCGAATATCTCCATTTTATGGTCTATCCCGCAGAGCGTCCGCGCA 729

DB 301 ThrGlyAspValIleHisMetSerProPheGlyLeuArg---AspGlyAlaHisVal 319
QY 730 GAACCCATGGGATATCCCGAGATAATTTCAACAACTAGATAGTATTTTTCATGAT 789
DB 320 GluHisThrSerTyrSerSerAspArgPheGlnGlnIleGluGlyTyrTyrProIleasp 339
QY 790 TTGGCAAGCGTCGAAAGCAAGCCCTCCAGTCAAGCGTAACCTTCTCATCATCACATCAC 849
DB 340 LeuAspThrAspTyrThrGlyAla---ProValSerArgAsnPheLeuGluThrProHis 358
QY 850 TTCACAGTTGGTGGGACTGGCTCCAAAACACTACTCTGTATGTTCAATCAGTAACGG 909
DB 359 ValThrValAlaTrpAsnTrpThrProLysSerGlyArgValCysThrLeuAlaLysTrp 378
QY 910 AAAGAGTGACTGAAATGTTGCTGCAACAGTAAATGGAGATACAGATTTATGCGCGT 969
DB 379 ArgGluIleAspGluMetLeuProMetAsnIle---GlySerTyrArgPheThrAlaLys 397
QY 970 GAACCTTCGCGCAACGTTTATCAGTAAATAGACTAGTTGATCCAAATCGCATCATATA 1029
DB 398 ThrIleSerAlaThrPheIleSerAsnThrSerGlnPheGluIleAsnArgIleArgLeu 417
QY 1030 GGCAATGTATTAACCGCAGCAGCAGCAGCAATCGACAGATATTTAGGACAAATAT 1089
DB 418 GlyAspCysAlaThrLysGluAlaAlaGluAlaIleAspArgIleTyrLysSerLysTyr 437
QY 1090 AATGACAGATCACGTCAGGTGTGACATGTACAATATTTCTTGCTCTCGGGGATTTAT 1149
DB 438 SerLysThrHisIleGlnThrGlyThrLeuGluThrTyrLeuAlaArgGlyGlyPheLeu 457
QY 1150 GTAGCATACAGCTGTTCTATCAATCCCTGGCTCATATGACCTCAGAGATTTGATG 1209
DB 458 IleAlaPheArgProMetIleSerAsnGluLeuAlaLysLeuIleAsnGluLeuAla 477
QY 1210 AGACACACAGGACCGATGAGATGCTCGACCTGGTAAACAATAAGCATCAATTTATAAG 1269
DB 478 ArgSerAsnArgThr-----ValValaspLeu-----SerAlaLeuLeuAsn 491
QY 1270 AAAATGCTACCTCATCTGTCACGATGCGCGAGATATTCGAATCGCAATCAACAA 1329
DB 492 ProSerGlyGluThrValGlnArgThrArgArgSerValProSerAsnGlnHisArg 511
QY 1330 ATA-----ACATTAGACGACACACAGCATTTAAA 1359
DB 512 SerArgArgSerThrIleGluGlyGlyIleGluThrValAsnAsnAlaSerLeuLys 531
QY 1360 TCGACATCGTGTTCATTAATTCGCGATGCTCCATTTCTTATCATATATACAAACCAT 1419
DB 532 ThrThrSerSerValGluPheAlaMetLeuGlnPheAlaTyrAspTyrIleGlnAlaHis 551
QY 1420 ATTAATGATATGTTTAGTATGCGACGCTGGTGGCAATTCGAGAATAGAGAACTT 1479
DB 552 ValAsnGluMetLeuSerArgIleAlaThrAlaTrpCysThrLeuGlnAsnArgGluHis 571
QY 1480 GTTTTATGGCAGAGGATAAAGATTAATCTTACGCGCTACAGGATGCAACATTAGA 1539
DB 572 ValLeuTrpThrGluThrLeuLysLeuAsnProGlyValValSerMetAlaLeuGlu 591
QY 1540 AGGAGATGCGTGCAGAGATGTTGGGGATGCTCCCTGCTGATCGACTGCTGCTATTA 1599
DB 592 ArgArgValSerAlaArgLeuGlyAspAlaValAlaValThrGlnCysValAsnIle 611
QY 1600 GATGCGGAATCCGTCACCTTTCGAAAATTCATGCGAGTTATTCACATCCACTTAAT 1659
DB 612 SerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrCys 631
QY 1660 TATACCGCAGCATGTTCTTATTTTCATATGCGAGTTATTCACATCCACTTAATCAT 1719
DB 632 TyrSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluTyrIleGluGln 651
QY 1720 CTCGCTGAAAACACAGCTGCTTCCACCGCTAGAGGCTGTAGAGCCATGCTCGCGTAAT 1779

```
Db 652 LeuGlyGluAsnAsnGluLeuLeuValGluArgLysLeuLeuGluProCysThrValAsn 671
Qy 1780 CATCGTAGATATTTCTGTTGGATCCGGTATCTTTATTTGAAACTATAATTTGTT 1839
Db 672 AsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyrVal 691
Qy 1840 AAGATGTAGACGTCGCGATATACAGTCTGACATTTGTCGAGCTTAATCTAAC 1899
Db 692 ArgLysValProLeuSerGluLeuLeuSerAlaTyrVal---IleLysSerThr 710
Qy 1900 CTGCTAGAAGATCGGAAATTTTCCTTTACACAAAGAGAGCTTCGCTGAT 1959
Db 711 LeuLeuGluAspArgGluPheLeuHis---SerSerTyrThrArgAlaGluLeuGluAsp 729
Qy 1960 GTTCGTGTATTTGATATGACAGATGACGTCGCGCAATCAATCAATCAATCAATTT 2019
Db 730 ThrGlyProPheAspTyrSerGluLeuGlnArgAsnGlnLeuHisAlaLeuLysPhe 749
Qy 2020 TATGACATAACAAGTAATAGAGTGTATCAAT----- 2055
Db 750 TyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMetAla 769
Qy 2056 -----TACGGGGGCTCGAGGAATTCGGCTGT----- 2082
Db 770 AsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGlyAla 789
Qy 2083 -----ATGCTATTACTAAAGATGCAACCAATAATATGCGCAAC 2127
Db 790 AlaSerAlaValIleSerThrValSerGlyValSerPheLeuAsnAsnProPheGly 809
Qy 2128 CAATTAGACGCGGATGAGTAAACAGATCAATCAATCAATCAAT----- 2172
Db 810 AlaLeuAlaValGlyLeuLeuLeuLeuAlaGlyIleValAlaAlaPheLeuAlaTyrArg 829
Qy 2173 -----AAAGCGATG-----ACATTAGCTTCA 2193
Db 830 TyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArgAsn 849
Qy 2194 CTACAAGACTATGCCAGATGAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAACA 2253
Db 850 LeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyLysAspSerAspProGlyVal 869
Qy 2254 GTTAACAATACCTTAATGCAACATAGAACAACTAAATGGCTAAA---ACTAATTA 2310
Db 870 AspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSerLeu 889
Qy 2311 GAATCAGCCATCAACCAAGCTAATACGGAATAAAGACGCTTTTGATAATGAACACCAAT 2370
Db 890 ValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyProAla 908
Qy 2371 TTAGTTGAAGCATAC-----AAAGCACTAAACAAAC-----ACTTTAGAACACGT 2415
Db 909 IleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgGlyProLysTyrGlnArg 928
Qy 2416 GCTACTACCTTGAGGTTTGTCACTCACTGCTTATATCAAT 2460
Db 929 LeuAsnAsnLeuAspSerGlyAspThrGluThrAsnLeuVal 943

RESULT 13
AAW50113
XX AAW50113 standard; Protein; 943 AA.
XX AC AAW50113;
XX DT 21-DEC-2001 (first entry)
XX DE Feline herpesvirus glycoprotein B recombinant antigen PFHVgB943.
XX FHV; glycoprotein B; PFHVgB943; antigen; immune status;
KW vaccination status; cat.
XX OS Feline herpesvirus.
XX
```

```
PN WO200166568-A2.
XX 13-SEP-2001.
XX 07-MAR-2001; 2001WO-US07251.
XX 09-MAR-2000; 2000US-0521738.
XX (HESK-) HESKA CORP.
XX (COLS ) UNIV COLORADO STATE RES FOUND.
XX Jensen WA, Lappin MR, Rosen DK, Andrews JS;
XX WPI; 2001-639000/73.
XX DR N-PSDB; AAH27060.
XX PT Determining immune status or vaccination status of an animal to e.g.
XX PT calicivirus comprises using a recombinant viral antigen -
XX PS Claim 4; Page 87-90; 132pp; English.
XX CC The present sequence is that of full-length feline herpesvirus
XX CC glycoprotein B recombinant antigen PFHVgB943. The recombinant
XX CC antigen, fused to an N-terminal His tag, was produced in Escherichia
XX CC coli cells transformed by a recombinant vector comprising nucleic
XX CC acid nPHVgB2829 (see AAH27060). PFHVgB943 is an example of a
XX CC recombinant infectious agent antigen that can be used in the method
XX CC of the invention to determine the immune status of an animal.
XX CC The method involves contacting a biological specimen of an animal
XX CC (cat, dog or horse) with a recombinant antigen, and detecting the
XX CC presence or absence of a complex between the recombinant antigen
XX CC and an antibody present in the sample. The method determines
XX CC whether the animal is protected against disease or should be
XX CC vaccinated. Recombinant antigens (see AAM50107-24), nucleic acids
XX CC encoding them (see AAH27054-71), methods of producing them, and
XX CC assay methods are provided.
XX SQ Sequence 943 AA;

Alignment Scores:
Pred. No.: 1.35e-128 Length: 943
Score: 1743.00 Matches: 377
Percent Similarity: 57.54% Conservative: 161
Best Local Similarity: 40.32% Mismatches: 267
Query Match: 30.77% Indels: 130
DB: 22 Gaps: 21

US-09-147-052-3 (1-3261) x AAM50113 (1-943)
Qy 7 TATTTTAGCGGAATTGCATATTT----- 30
Db 22 TyrPheArgGlnArgCysPhePheProSerLeuLeuGlyIleAlaAlaThrGlySerArg 41
Qy 31 -----TTCTTATAGTT 42
Db 42 HisGlyAsnGlySerSerGlyLeuThrArgLeuAlaArgTyrValSerPheIleTrpIle 61
Qy 43 ATTTCTATAT-----GGTACCAACTCATCTCCG 69
Db 62 ValLeuPheLeuValGlyProArgProValGluGlyGlnSerGlySerThrSerGluGln 81
Qy 70 AGTACCCAAATGTGACATCAAGAGAGTT----- 99
Db 82 ProArgArgThrValAlaThrProGluValGlyValHisHisGlnAsnGlnLeuGlnIle 101
Qy 100 -----GTTTCGAGCGCTCAGTTGCTGAGGAGAG----- 129
Db 102 ProProfileCysArgTyrGluGluAlaLeuArgAlaSerGlnIleGluAlaAsnGlyPro 121
Qy 130 TCTACGTTTATCTTTGTCCTCCACACGCTGGGTCAACCGTGATCCGCTAGAACCGCG 189
Db 122 SerThrPheTyrMetCysProProSerGlySerThrValValArgLeuGluProPro 141
```

[illegible]

Db	492	ProSerGlyClnThrValGlnArgThrArgSerValProSerAsnGlnHisArg	511
Qy	1330	ATA-----ACATTAGGACACACACAGCTATTAAA	1359
Db	512	SerArgArgSerThrIleGluGlyIleGluThrValAsnAsnAlaSerLeuLeuLys	531
Qy	1360	TCGCATCGCTGTTCAATTCGCCATGCTTCTTATGATCATATACAAACCCAT	1419
Db	532	ThrThrSerValGluPheAlaMetLeuGlnPheAlaIleValThrSerPheGlnAlaHis	551
Qy	1420	ATTAAATCATATGTTTACTAGATTGCGACAGCTTGGTGGCAATTCAGAAATGCAGAAATAGAGAACTT	1479
Db	552	ValAsnGlnMetLeuSerArgIleAlaThrAlaIleValThrSerPheGlnAsnArgGluHis	571
Qy	1480	GTTTTATTGGCAGCAGAGGATAAAGATTAATCTTAGCGCTACAGCGAGTCCACATAGGA	1539
Db	572	ValLeuTrpThrGluThrLeuLysLeuAsnProGlyGlyValIleSerMetAlaLeuGlu	591
Qy	1540	AGGAGATGGCTGCAAGAGTGTGGGGAGTGCCTGCTGCTATTCAGAGTGCCTACTGCTATA	1599
Db	592	ArgArgValSerAlaArgLeuLeuGlyAspAlaValAlaValThrGlnCysValAsnIle	611
Qy	1600	GATCGGGAATCCGTCATCTTGCATAATTCATCGGAGTTATCACATCCACTAATACAGT	1659
Db	612	SerSerGlyHisValTyrIleGlnAsnSerMetArgValThrGlySerSerThrThrCys	631
Qy	1660	TATAGCCGACCATTTGGTCTATTTCATATGGAGAAACCAAGAAACATACAGGACAA	1719
Db	632	TyrSerArgProLeuValSerPheArgAlaLeuAsnAspSerGluThrIleGluGlyGln	651
Qy	1720	CTCGGTGAAAAACAGAGTTCCTCCAACTAGAGCTAGAGCTCTAGAGCCATCTCGGCTAAT	1779
Db	652	LeuGlyGluAsnAsnGluLeuValGluArgLysLeuIleGluProCysThrValAsn	671
Qy	1780	CATCTAGATATTTCTGTTGGATCGGTTATCTTATTTCGAAACATATAATTTGTT	1839
Db	672	AsnLysArgTyrPheLysPheGlyAlaAspTyrValTyrPheGluAspTyrAlaTyrVal	691
Qy	1840	AAGATGGTAGCGCTGCCGATATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAAC	1899
Db	692	ArgLysValProLeuSerGluIleGluLeuIleSerAlaTyrValIleLysSerThr	710
Qy	1900	CTGCTAGAGATCGGGAATTTTGCCTTTATCCGTTTACACAAAGAAAGAGTTGCGTGAT	1959
Db	711	LeuLeuGluAspArgGluPheLeuHis---SerSerTyrThrArgAlaGluLeuGluAsp	729
Qy	1960	GTTGTGCTATTGATTTATGCAGATAGCTCGCCCAATCAACTACATGAACTTAATTT	2019
Db	730	ThrGlyProPheAspTyrSerGluIleGlnArgArgAsnGlnLeuHisAlaLeuLysPhe	749
Qy	2020	TATGACATAACAAAGTAAATAGAACTGGATCAAAAT-----	2055
Db	750	TyrAspIleAspSerIleValArgValAspAsnAsnLeuValIleMetArgGlyMetAla	769
Qy	2056	-----TACCGGGGCTCGAGGAATTCGCTGT-----	2082
Db	770	AsnPhePheGlnGlyLeuGlyAspValGlyAlaGlyPheGlyLysValValLeuGlyAla	789
Qy	2083	-----ATGCTCTATTACTAAAAAGATGCAAAACCCAAATAATATGCGCAAC	2127
Db	790	AlaSerAlaValIleSerThrValSerGlyValSerSerPheLeuAsnAsnProPheGly	809
Qy	2128	CAATTAGACGACGGGAATGAGTTACAGATCTAATCATGCT-----	2172
Db	810	AlaLeuAlaValGlyLeuLeuIleLeuAlaGlyIleValAlaAlaPheLeuAlaTyrArg	829
Qy	2173	-----AAAGCGATG-----ACATTAGCTTCA	2193
Db	830	TyrIleSerArgLeuArgAlaAsnProMetLysAlaLeuTyrProValThrThrArgAsn	849
Qy	2194	CTACAAGACATTCGAAGATTGAAGTAGTTTATCATCTGCTTATAGTGAAGCTGAAACA	2253

Db	850	LeuLysGlnThrAlaLysSerProAlaSerThrAlaGlyClyAspSerAspProGlyVal	869
Qy	2254	GTTTAAACAATAACCTTAAATGCAACATTAGAACAACTAAATAATGGCTAAA---ACTAATTTTA	2310
Db	870	AspAspPheAspGluGluLysLeuMetGlnAlaArgGluMetIleLysTyrMetSerLeu	889
Qy	2311	GAATGAGCCATCAACCAAGCTAATACGGATAAAACGACATTTTGATTAATGAACACCCAAAT	2370
Db	890	ValSerAlaMetGluGln---GlnGluHisLysAlaMetLysLysAsnLysGlyProAla	908
Qy	2371	TTTAGTTGAACCATAC-----AAAGCACTAAAAACC-----ACTTTAGACAAACGT	2415
Db	909	IleLeuThrSerHisLeuThrAsnMetAlaLeuArgArgGlyProLysTyrGlnArg	928
Qy	2416	GCTACTAAACCTTGAAGGTTTGTCATCAACATGCTTATAATCAAAATT	2460
Db	929	LeuAsnAsnLeuAspSerGlyAspAspThrGluThrAsnLeuVal	943
RESULT	14		
AAR92746	ID	AAR92746 standard; protein; 891 AA.	
XX	AC	AAR92746;	
XX	XX	11-MAY-1996 (first entry)	
DT	XX	B virus gB glycoprotein.	
DE	XX	Herpes simian monkey B virus gB glycoprotein; UL27; immunoassay;	
KW	KW	diagnosis; herpes B virus.	
XX	XX	Herpes simian monkey B virus.	
OS	XX	US5487969-A.	
PN	XX	30-JAN-1996.	
PD	XX	01-APR-1993; 93US-0042747.	
PF	XX	01-APR-1993; 93US-0042747.	
PR	XX	(SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.	
PA	XX	Black D, Eberle R, Hilliard J, Scinicariello F;	
PI	XX	WPI; 1996-105220/11.	
DR	DR	N-PSDB; AAT16474.	
PT	PT	Detection of herpes B virus by PCR amplification of sample DNA - to	
PT	PT	detect a specific herpes simian monkey B virus DNA segment.	
XX	XX	Claim 1; Column 19-24; 22pp; English.	
PS	XX	The herpes simian monkey B virus proteins, such as the gB	
CC	CC	glycoprotein (UL27), have immense potential use in the development	
CC	CC	of serological immunoassays, which can specifically detect virus	
CC	CC	antigens and/or antibodies to B virus. One approach is to	
CC	CC	synthesize peptides which, based on the properties of the predicted	
CC	CC	protein sequence, are likely to be immunologically active. Such	
CC	CC	peptides can be used as substrate antigens in immunoassays to detect	
CC	CC	serum antibodies which recognize this specific peptide sequence.	
CC	CC	Synthetic peptides may also be used to produce antibodies against	
CC	CC	specific regions of the gB glycoprotein which are unique to one	
CC	CC	virus. These can then be used to develop virus-specific	
CC	CC	immunoassays for differentiation of B virus from other primate	
CC	CC	alpha-herpes virus and for identification of antibodies directed	
CC	CC	against B virus in primate serum samples.	
XX	XX	Sequence 891 AA;	
SQ	SQ	Alignment Scores:	
		Pred. No.:	2.51e-128
		Score:	1739.50
		Length:	891
		Matches:	332

Percent Similarity:	64.16%	Conservative:	103
Best Local Similarity:	48.97%	Mismatches:	216
Query Match:	30.71%	Indels:	27
DB:	17	Gaps:	3

US-09-147-052-3 (1-3261) x AAR92746 (1-891)

Qy	52	GGTACGAACATCTCCGAGTACCMAAATGTGCATCAAGAAGAAGTTGTTTCGAGCGTC	111
Db		::: :::
Db	72	GlyThrAsnAlaSerValGIuAlaGlyHisAlaThrLeuArgGIuAsnLeuArgAspIle	91
Qy	112	CAGTTGTCAGGAAGAGTCTACGTTTATCTTTGTCGCCACACAGTGCGTCAACCGTG	171
Db		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db	92	LysAlaLeuAspGlyAspAlaThrPheTyValCysProProThrGlyAlaThrVal	111
Qy	172	ATCCGCTAGAACCGCGCGAAAATGTCGCCGAACCTAGAAAGACCCAGTAGTGGGTGAA	231
Db		:::::	::: :::: :::: ::::
Db	112	ValGlnPheGIuGlnProArgProCysProArgAlaProHisGlyGlnAsnTyThrGlu	131
Qy	232	GGATCCGCGATATTATTAAAGAGAATATCAGTCCCATATAAATTTAAAGTGACGCTTAT	291
Db	
Db	132	GlyIleAlaValIlePheLysGIuAsnIleAlaProTyrLysPheLysAlaThrMetYr	151
Qy	292	TATAAAATATCATCTACAGCAGCACATGCGGGGAGCAGATATACAGACATCACTAAT	351
Db	:::
Db	152	TyrLysaspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGly	171
Qy	352	CGATATACAGATAGGACGCCGTTTCATTGACAGAGATCAGGATCTAATTCGAGCGCAA	411
Db		:::	:::: :::: ::::
Db	172	IlePheGIuAspArgAlaProValProPheGIuGluValIleAspLysIleAsnAlaArg	191
Qy	412	GGAAGATGCTCATCTAAAGCAAGATACCTTAGAAACAATGTATATGTTGAAGCGTTTGAC	471
Db	
Db	192	GlyValCysargSerThrAlaLysTyValArgAsnAsnMetGluSerThrAlaPheHis	211
Qy	472	AGGGATCGGGAGAAAAACAAGTACTTCTAAAACCATCAAAATTCACACGCCCGATCT	531
Db		::::
Db	212	ArgAspaspGluSerAspMetLysLeuLysProAlaLysAlaAlaThrArgThrSer	231
Qy	532	AGGCATGGSCACGACATATGAGACGTATACCGGTGGGGATCACCATCGATATATCGA	591
Db	
Db	232	ArgGlyTrpHisThrThrAspLeuLysTyAsnProSerArgIleGIuAlaPheHisArg	251
Qy	592	ACGGGAACCTCCGTCATTTGATAGTAGAGAAATGGATCCGCTCTGTGTTCCGTAT	651
Db	
Db	252	TyrGlyThrThrValAsnCysIleValGIuGIuValGIuAlaArgSerValTyrProTyr	271
Qy	652	TCATATTTTGCANTGGCCATGGGACATCGCGAACATATCTCCATTTTATGTCATACC	711
Db		::::
Db	272	AspGluPheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg	291
Qy	712	CCACCAGAGCTCCCGCAGAACCCATCGGATATCCCCAGGATATTTTCAACAACACTAGAT	771
Db		
Db	292	--- AspGlyAlaHisAlaGIuHisThrAlaTyrAlaAlaAspArgPheArgGlnValAsp	310
Qy	772	AGCTATTTTCAATGGATTGGCAACGCTCGAAAAGCAAGCCCTCCAGTCAACGCGTAAC	831
Db	
Db	311	GlyTyrThrGluArgAspLeuSerThrGlyArgArgAlaSerThrProAlaThrArgAsn	330
Qy	832	TTTCTCATCATCATCACACTTCACAGTTGGGTGGGACTGGGCTCCAAAAACTACTCGTGA	891
Db		
Db	331	LeuLeuThrThrProLysPheThrValGlyTrpAspTrpAlaProLysArgProSerVal	350
Qy	892	TGTTCAATGACTAAGTGGAAAGAGGTGACTGAAATGTTGGTGGCAACAGTTAATGGAGA	951
Db	
Db	351	CysThrLeuThrTystrpGIuValAspGIuMetLeuArgAlaGIuTyrGlyProSer	370
Qy	952	TACAGATTTATGGCCCGTGAACCTTCGCGAACGTTTATCATGATACGACTGAGTTTGAT	1011
Db		::: ::::
Db	371	PheArgPheSerSerSerAlaLeuSerThrThrPheThrThrAsnArgThrGluTyrAla	390
Qy	1012	CCAATTCGCATCATATTAGGACAAATGATTAAACGCGGAGCAGCAAGCAGCAATTCGACAG	1071
Db	

Db	391	LeuSerArgValAspLeuGlyAspCysValGlyArgGluAlaAargGluAlaValAspArg	410
QY	1072	ATATTTAGGACAAATATAATGACAGCTACAGCTCAAGGTTGGACATGTCACAAATATTCCTG	1131
Db	411	IllePheLeuArgArgTyrAsnGlyThrHisValLysValGlyGlnValGlnTyrTyrLeu	430
QY	1132	GCTCTCGGGGATTTATGTAGCATATACGCTGTTCATCCAATCCCTGGCTCATATG	1191
Db	431	AlaThrAlaGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnGlyLeuValGluLeu	450
QY	1192	TACCTCAGAGAAATGATGAGACAGACAGCCGATGACATGCTCGACCTGGTAAACAAT	1251
Db	451	TyrValArgGluLeuLeuArgGluGlnGluGlyArgPro	463
QY	1252	AAGCATGCAATTTATAAGAAAAATGCTACCTCATCTCAGATGCGGGCAGATATCGA	1311
Db	464	-----GlyAspAlaAla	467
QY	1312	ANTGCACCAATAGAAAAATAACATTAGAGCACACACAGCTATTAAATCGCATCGTCT	1371
Db	468	AlaThrProLysProSerAlaAspProAspValGluArgIleLysThrThrSerSer	487
QY	1372	GTTCAATTCGCCATGCTCCAAATTTCTTTATGATCATATACAACCCOATTAATGATATG	1431
Db	488	ValGluPheAlaArgLeuGlnPheThrTyrAspHisIleGlnArgHisValAsnAspMet	507
QY	1432	TTTATGATAGGATGCCACAGCTTGGTGGAAATCCAGAATTCCAGAACTGTTGTTATGGCAC	1491
Db	508	LeuGlyArgIleAlaIleAlaIleTyrCysGluLeuGlnAsnHisGluLeuThrLeuTyrPsn	527
QY	1492	GAAGGATAAAGATTAACTAGCGCTACAGCGAGTCAACATAGGAAGAGAGTGGCT	1551
Db	528	GluAlaArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgArgValSer	547
QY	1552	GCAAGATGTTGGGGATGTCGGTGCRTATCGAGCTGCATGCTATAGATGCGGAATCC	1611
Db	548	AlaArgMetLeuGlyAspValMetalValSerThrCysValProValThrProAspAsn	567
QY	1612	GTCACCTTTGCAAAATCTATGTCGAGTTATCACATCCACTAATACATGTTATATAGCCGACCA	1671
Db	568	ValIleMetGlnAsnSerMetArgValProAlaArgProGlyThrCysTyrSerArgPro	587
QY	1672	TGTTGTTATTTTCATATGGAGAAACCAAGGAACATACAGGACAACTCGGTGAAC	1731
Db	588	LeuValSerPheArgTyrGluGluGlyGlyProLeuValGluGlyGlnLeuGlyGluAsp	607
QY	1732	AACGATGTCCTCAACGCTAGAGCGCTAGAGCGATGCTCGGCTAATCATCTAGATAT	1791
Db	608	AsnGluIleArgLeuGluArgAspAlaLeuGluProCysThrValGlyHisArgArgTyr	627
QY	1792	TTTCTGTTTGGATCCGGTTATGCTTTATTTGAAACCTATAATTTGTTAGATCGGTAGAC	1851
Db	628	PheThrPheGlyAlaGlyTyrValTyrPheGluAspTyrAlaTyrSerHisGlnLeuGly	647
QY	1852	GCTGCGGATATACAGATTGCTAGCACATTTGTCGAGCTTAATCTAACCCCTAGTAGAAGAT	1911
Db	648	ArgAlaAspValThrThrValSerThrPheIleAsnLeuAsnLeuThrMetLeuGluAsp	667
QY	1912	CGGGAATTTTGCCTTTATCCGTTTACACAAAGAAGATGTCGTGATGTTGGTGTATTG	1971
Db	668	HisGluPheValProLeuGluValTyrThrArgGlnIleLysAspSerGlyLeuLeu	687
QY	1972	GATTATGACAGATGACGTCGGCGCAATCACTACATGAACTTAATTTTATGACATAAAC	2031
Db	688	AspTyrThrGluValGlnArgAsnGlnLeuHisAlaLeuArgPheAlaAspIleAsp	707
QY	2032	AAAGTAATAGAGTGCATACAAAT-----TACGGGGCTCGAGGAATTC	2076
Db	708	ThrValIleLysAlaAspAlaHisAlaProLeuPheAlaGlyLeuTyrSerPhe	725

RESULT 15

AAW70293

ID AAW70293 standard: Protein; 891 AA.

Db 152 TyrLysAspValThrValSerGlnValTrpPheGlyHisArgTyrSerGlnPheMetGly 171
 Qy 352 CGATATACAGTAGAGCCCGTTTCATTAAGAGATACAGGATCTAATACGAGCGGAAA 411
 Db 172 IlePheGluAspArgAlaProValProPheGluGluValIleAspLysIleAsnAlaArg 191
 Qy 412 GGAAGATGCTCANTAAAGCAAGATACCTTAGAACAATGTATATGTTGAAGCGTTTGAC 471
 Db 192 GlyValCysArgSerThrAlaLysTyrValArgAsnAsnMetGluSerThrAlaPheHis 211
 Qy 472 AGGATCGGGAGAAAAACAAGTACTTCAAAACCATCAAAATCAACACGCGCGAATCT 531
 Db 212 ArgAspAspAspGluSerAspMetLysLeuLysProAlaLysAlaAlaThrArgThrSer 231
 Qy 532 AGGCATCGGCACGACGACTAATAGACGCTATACCGGTGGGATACACCATGATATATCGA 591
 Db 232 ArgGlyThrHisThrAspLeuLysTyrAsnProSerArgIleGluAlaPheHisArg 251
 Qy 592 ACGGAACCTCCGTCATTTGATAGTAGAGGAATGGATGCCGCTGTGTTCCGTAT 651
 Db 252 TyrGlyThrThrValAsnCysIleValGluGluValGluAlaArgSerValTyrProTyr 271
 Qy 652 TCATATTTGCAATGGCAATGGCCACATCGCGAACATATCTCCATTTTATGGTCTATCC 711
 Db 272 AspGluPheValLeuAlaThrGlyAspPheValTyrMetSerProPheTyrGlyTyrArg 291
 Qy 712 CCACGAGGCTGCGCGAGACCCATCGGATATCCCGAGGATATTTCAACAACACTAGAT 771
 Db 292 ---AspGlyAlaHisAlaHisThrAlaTyrAlaAlaAspArgPheArgGlnValAsp 310
 Qy 772 AGCTATTTTCAATGGATTTGGACAGCGTCGAAAGCAAGCCTTCAGTCACGCGTAAC 831
 Db 311 GlyTyrTyrGluArgAspLeuSerThrGlyArgAlaSerThrProAlaThrArgAsn 330
 Qy 832 TTTCTCATCATCACATCACAGTTGGGTGGGACTGGGCTCCAAAACACTACTCGGTGA 891
 Db 331 LeuLeuThrThrProLysPheThrValGlyTrpAspTrpAlaProLysArgProSerVal 350
 Qy 892 TGTTCAATGACTAAGTGAAGAGGTGACTGAATGTGGGTGCGTCAACAGTTAATGGAGA 951
 Db 351 CysThrLeuThrLysTrpGlnGluValAspGluMetLeuArgAlaGluTyrGlyProSer 370
 Qy 952 TACAGATTTATGGCCCGTGAACCTTCGGCAACGTTTATCAGTATACGACTGAGTTTGT 1011
 Db 371 PheArgPheSerSerAlaLeuSerThrThrPheThrThrAsnArgThrGluTyrAla 390
 Qy 1012 CCAATCGCATATATTAAGCAATGATTAAGCGGAGCAGCAAGCAGCAATCGAGCAG 1071
 Db 391 LeuSerArgValAspLeuGlyAspCysValGlyArgGluAlaArgGluAlaValAspArg 410
 Qy 1072 ATATTTAGGCAAAAATATATGACAGTCACGTCAGGTTCAGCATGTGACATATTTCTTG 1131
 Db 411 IlePheLeuArgArgTyrAsnGlyThrHisValLysValGlyGlnValGlnTyrTyrLeu 430
 Qy 1132 GCTCTCGGGGATTTATGTAGCATATCAGCTGCTTCTATCCAAATCCCTGGCTCATATG 1191
 Db 431 AlaThrAlaGlyPheLeuIleAlaTyrGlnProLeuLeuSerAsnGlyLeuValGluLeu 450
 Qy 1192 TACCTCAGAGATTTATGTAGCAGACACAGGACCGCATGAGATGCTCGACCTGGTAAACAAT 1251
 Db 451 TyrValArgGluLeuLeuArgGluGlnGlyArgPro----- 463
 Qy 1252 AAGCATGCAATTTATAGAAAAATGCTACCTCATTTGTCAGATTTGGCGGAGATATTCGA 1311
 Db 464 -----GlyAspAlaAla 467
 Qy 1312 AATCCACCAATAGAAAAATAACATTAGACGACACACGACTTATTAATCGACATCTCT 1371
 Db 468 AlaThrProLysProSerAlaAspProProAspValGluArgIleLysThrThrSerSer 487
 Qy 1372 GTTCAATTCGCGCATGCTCCATTTCTTATCATATACAAACCCCATATTAATGATATG 1431
 Db 488 ValGluPheAlaArgLeuGlnPheThrTyrAspHisIleGlnArgHisValAsnAspMet 507

Qy 1432 TTAGTAGTAGATTCGCACAGCTTGGTGCGAATTCGCAATATGCAATATAGAACACTGTTTATGCGAC 1491
 Db 508 LeuGlyArgIleAlaIleAlaTrpCysGluLeuGlnAsnHisGluLeuThrLeuTrpAsn 527
 Qy 1492 GAAGGGATAAAGATTATCTAGCGCTACAGCGAGTGCACACATTAGGAAGGAGAGTGGCT 1551
 Db 528 GluAlaArgLysLeuAsnProAsnAlaIleAlaSerAlaThrValGlyArgValSer 547
 Qy 1552 GCAAGATGTTGGGGAGTGCCTGCTGTATCGAGCTGCACTGCTATAGATGCGGAATCC 1611
 Db 548 AlaArgMetLeuGlyAspValMetAlaValSerThrCysValProValThrProAspAsn 567
 Qy 1612 GTCACCTTGCAAAATTCATGCGAGTTATCACATCCACTAATACATGTTATAGCCGACCA 1671
 Db 568 ValIleMetGlnAsnSerMetArgValProAlaArgProGlyThrCysTyrSerArgPro 587
 Qy 1672 TTGGTTCTATTTTCATATGGAAGAACCAAGAACATACAGGACAACTCGGTGAACAC 1731
 Db 588 LeuValSerPheArgTyrGluGlyGlyProLeuValGluGlyGlnLeuGlyGluAsp 607
 Qy 1732 AACGAGTTGCTTCCACGCTAGAGGCTGTAGAGCATGCTCGCTAATCATCTAGATAT 1791
 Db 608 AsnGluIleArgLeuGluArgAspAlaLeuGluProCysThrValGlyHisArgArgTyr 627
 Qy 1792 TTTCTGTTGGATCCGTTATGCTTTATTTGAAACTATATATTTGTTAAGATGCTAGAC 1851
 Db 628 PheThrPheGlyAlaGlyTyrValTyrPheGluAspTyrAlaTyrSerHisGlnLeuGly 647
 Qy 1852 GCTGCCATATACAGATGCTAGCACATTTGTCGAGCTTATCTAACCTGCTAGAGAT 1911
 Db 648 ArgAlaAspValThrThrValSerThrPheIleAsnLeuAsnLeuThrMetLeuGluAsp 667
 Qy 1912 CGGAAATTTTGCCTTTATCGTTTACACAAAAGAGTTGCTGATGTTGGTGTATTG 1971
 Db 668 HisGluPheValProLeuGluValTyrThrArgGlnGluIleLysAspSerGlyLeuLeu 687
 Qy 1972 GATTATGCAGAGTAGCTCGCGCAATCAACTACATGAACCTTAAATTTTATGACATAAC 2031
 Db 688 AspTyrThrGluValGlnArgArgAsnGlnLeuHisAlaLeuArgPheAlaAspIleAsp 707
 Qy 2032 AAGTAATAGAGTGGATACAAAT-----TACGGGGGCTGCGAGGAATTC 2076
 Db 708 ThrValIleLysAlaAspAlaHisAlaProLeuPheAlaGlyLeuTyrSerPhe 725

Search completed: October 8, 2003, 17:53:37
 Job time : 270.496 secs

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:47:17 ; Search time 48.58 Seconds
(without alignments)
1399.486 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSITKDPNNGQQLA.....SSNNADKIPGVRPQTF 393

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	60.1	271	2 Q49464	Q49464 mycoplasma
2	837	41.9	671	2 Q9L8D6	Q9L8D6 mycoplasma
3	831.5	41.6	702	2 Q49499	Q49499 mycoplasma
4	826	41.3	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	40.3	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	38.1	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	37.7	702	2 Q49498	Q49498 mycoplasma
8	746	37.3	645	2 Q9Z1D1	Q9Z1D1 mycoplasma
9	745	37.3	644	2 Q9ZHR9	Q9ZHR9 mycoplasma
10	744	37.2	650	2 Q49495	Q49495 mycoplasma
11	743	37.2	649	2 Q49497	Q49497 mycoplasma
12	713	35.7	647	2 Q49468	Q49468 mycoplasma
13	704	35.2	656	2 Q9KH15	Q9KH15 mycoplasma
14	690.5	34.6	703	2 Q05122	Q05122 mycoplasma
15	619.5	31.0	419	2 Q9L8D5	Q9L8D5 mycoplasma
16	575.5	28.8	386	2 Q49500	Q49500 mycoplasma

17	557	27.9	367	2 Q9XCG7	Q9XCG7 mycoplasma
18	189.5	9.5	320	2 Q49496	Q49496 mycoplasma
19	175	8.8	6713	16 Q99U54	Q99U54 staphylococ
20	175	8.8	6713	16 Q931R6	Q931R6 staphylococ
21	173	8.7	1302	2 Q49547	Q49547 mycoplasma
22	164.5	8.2	2481	16 Q99Q96	Q99Q96 staphylococ
23	163.5	8.2	3890	16 Q99U53	Q99U53 staphylococ
24	160.5	8.0	433	16 Q97TE6	Q97TE6 clostridium
25	160.5	8.0	682	2 Q48674	Q48674 lactococcus
26	159	8.0	933	2 Q53653	Q53653 staphylococ
27	158.5	7.9	1344	2 Q49545	Q49545 mycoplasma
28	156	7.8	807	5 Q96262	Q96262 plasmodium
29	155	7.8	1072	16 Q9CF64	Q9CF64 lactococcus
30	155	7.8	4688	16 Q9PQ08	Q9PQ08 ureaplasma
31	154.5	7.7	1365	2 Q49525	Q49525 mycoplasma
32	153.5	7.7	1795	2 Q9LCJ9	Q9LCJ9 staphylococ
33	153.5	7.7	2478	2 Q9RL69	Q9RL69 staphylococ
34	151.5	7.6	1237	16 Q9ZK57	Q9ZK57 helicobacte
35	151.5	7.6	2478	2 Q9LCH2	Q9LCH2 staphylococ
36	151	7.6	624	2 Q49548	Q49548 mycoplasma
37	150.5	7.5	604	5 Q26021	Q26021 plasmodium
38	150.5	7.5	1051	2 Q49524	Q49524 mycoplasma
39	149.5	7.5	962	2 Q49546	Q49546 mycoplasma
40	148	7.4	716	2 Q49526	Q49526 mycoplasma
41	147.5	7.4	655	5 Q26109	Q26109 plasmodium
42	145	7.3	2399	16 Q9ZK89	Q9ZK89 helicobacte
43	144	7.2	1107	2 Q9F2D8	Q9F2D8 salmonella
44	143.5	7.2	751	3 Q94358	Q94358 schizosacch
45	143	7.2	1034	3 Q9Y7E9	Q9Y7E9 candida gla

ALIGNMENTS

RESULT 1
Q49464 ID Q49464 PRELIMINARY; PRT; 271 AA.
AC Q49464;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)
DE TM-1 (FRAGMENT).
GN TM-1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94025993; PubMed=8212828;
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,
RA Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.;
RT "Cloning and DNA sequence of a 29 kilodalton polypeptide gene of
RT Mycoplasma gallisepticum as a possible protective antigen.";
RL Vaccine 11:1061-1066(1993).
DR EMBL; S65869; AAB28343.2; -.
FT NON_TER 271
SQ SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;

Query Match 60.1%; Score 1200; DB 2; Length 271;
Best Local Similarity 97.1%; Pred. No. 6.2e-52;
Matches 238; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 CMSITKDPNNGQQLAARNELDLINAKAMTASLQDYAKIESLSSAYSEATVN 60
DB 27 CMSITKDPNNGQQLAARNELDLINAKAMTASLQDYAKIESLSSAYSEATVN 86
QY 61 NNUNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVAYKALKTTLEQATNLEGLS 120
DB 87 NNUNATLEOLKMAKTNLESAINOANTDKTTFDNEHNPVAYKALKTTLEQATNLEGLA 146
QY 121 STAYNOIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQKTN 180

```
Db 147 STAYNOIRNLVDLYNNASSLIKTLDPLNGMLDSNEITTYNRNINNTLSINEQKTN 206
QY 181 ADALSNFIKKVIONNEQSVFGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVNGDE 240
Db 207 ADALSNFIKKVIONNEQSVFGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVNGDE 266
QY 241 PSSRI 245
Db 267 PSSRI 271

RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
AC Q9L8D6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PGMA-LIKE PROTEIN 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel PGMA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP210770; AAF29524.1; -.
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;

Query Match 41.9%; Score 837; DB 2; Length 671;
Best Local Similarity 43.4%; Pred. No. 1.1e-33;
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;

QY 8 DAPN-----NGOTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTNN 61
Db 55 DTNPNGDGMENSAQAQAAAKKELSDLLATQSSNLAKYADYTNQNTLTAAAYTTAKSTSD 114
QY 62 NLNATLEQKMAKTNLESAINQANTDKTTDFNEHPNVEAYKALKTTLEQRATNLEGLSS 121
Db 115 NTSVTLQVKSATSTLQATDAITRAASSKTSFDEKNPELIKAYALKETLKKEETVLSGLTD 174
QY 122 TAYNOIRNLVDLYNNASSLIKTLDPLNGMLDSNEITTYNRNINNTLSINEQKTNA 181
Db 175 SNFATIKTNLTALYQSGKDFVKATLDPVSGNA-PQADITKADKRDADAVSKLETWKTN 233
QY 182 DALSNFIKKVIONNEQSVFGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARRTVN 237
Db 234 NTLATSFVKELVKNLKTGIDT-TNNQEQPCNTSFVGSYNATNNNEIPNWNFAQRKVWT 292
QY 238 GDPESSRILA---NTNSITDWSIYSLAGTNTKYQSFNSYGPSTGILYFPYKLVK 293
Db 293 SDNGRTSLISSTSDNSSTLQAEVSWIYSLGAGTKYSLTFNYGPGSTGILYFPYKLVK 352
QY 294 ANNVGLQYKLNNGVQVEFA---TSTSA-----NNTT-----ANPTPAVD 331
Db 353 ENNVALQYTLNLSGAGEVNPATPVKYSVSDSGDSNNQTESAAETMPVTSIDLNPAPTVS 412
QY 332 EIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 383
Db 413 DINIAKLTLSNLAFSGNTIEFSVPTTAKETSKVAPMIGNIYLSNENAEKV 462

RESULT 3
Q49499 PRELIMINARY; PRT; 702 AA.
ID Q49499
AC Q49499;
```

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN PMGAI.4 PROTEIN PRECURSOR.
DE PMGAI.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
DR EMBL: L28424; AAA62418.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;

Query Match 41.6%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 2.1e-33;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 8 DAPNNG-----QTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTNN 61
Db 68 NTFPGNGGGTDNAAQAAAKKELSDLLATQNSNLSTADYANTIONTLTAAYTTAKSTSD 127
QY 62 NLNATLEQKMAKTNLESAINQANTDKTTDFNEHPNVEAYKALKTTLEQRATNLEGLSS 121
Db 128 NTSATLEQVKSATSTLQATDAITRAASSKTSFDEKNPELIKAYALKETLKWRNLSGLTD 187
QY 122 TAYNOIRNLVDLYNNASSLIKTLDPLNGMLDSNEITTYNRNINNTLSINEQKTNA 181
Db 188 SNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTN 246
QY 182 DALSNFIKKVIONNEQSVFGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARRTV 235
Db 247 TVLATSFVKELVKNLKTGIDT-TNNQEQPCNTSFVGSYVDVTTCSDNAPNWSFAQRKV 305
QY 236 WNGD-----BPSSRILATNSITDWSIYSLAGTNTKYQSFNSYGPSTGILYFPYKLVK 290
Db 306 WTSNTDILSQPAPAEQENQOSAPDVSWIYNTLGMGAKYSLTFNYGPGSTGILYFPYKLVN 365
QY 291 AADANNVGLQYKLNNGVQVEFATS-----TSANN-----TTANPTPAVD 331
Db 366 SSQSDKVALLEYKLNESAVKTDIDFSPQTSVPASDATRENNRSTAAAPAGGSTEINPAPTL 425
QY 332 EIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADK 382
Db 426 DIKIAKLTLSNLAFSGNTIEFSVPTTAKETSKVAPMIGNIYLSNENADK 477

RESULT 4
Q9XCG8 PRELIMINARY; PRT; 632 AA.
ID Q9XCG8
AC Q9XCG8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VLHAI PRECURSOR (FRAGMENT).
GN VLHAI.
OS Mycoplasma imitans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=29560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4229;
```

RA	MEDLINE=99392472; PubMed=10463176;	
RT	Markham P.F., Duffy M.F., Glew M.D., Browning G.F.;	
RL	"A gene family in Mycoplasma imitans closely related to the pmga	
RL	family of Mycoplasma gallisepticum.";	
RL	Microbiology 145:2095-2103(1999).	
DR	EMBL: AF141940; RAD39483.1; -	
KW	Signal	
FT	NON_TER 1 1	
FT	SIGNAL <1 17	POTENTIAL.
FT	CHAIN 18 632	VLHAI.
SQ	SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;	
Query Match	41.3%;	Score 826; DB 2; Length 632;
Best Local Similarity	45.3%;	Pred. No. 3.4e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps		
QY	7	KDANPNNGQT-----QLEAARMELDLNKAAMTLASLDQYAKIEASLSAYSAE 57
DB	34	KPNPDTGNTNPGDGTDAKQADAKASLATLGSQSTNVALYEDYAKIDTILSSAYSAQ 93
QY	58	TVNNNLNATLEQLKMAKTNLIESAINQANTDKTTFDNEHNPVLEYAKLTLEQRATNLE 117
DB	94	TTANNANATLKVDNDAKAAQLQTAISDAANAQTEFDKANLGSLYSAVAKLETILKSETNLD 153
QY	118	GLSTAYNQIRNLVDLYNKASSLTKTLDPLNGTLLDSNETTANKNINTLSTINEQ 177
DB	154	GLSQDNYSAIKANLVSLYKAKDFTVTLDLPTSGMIPKVDITTSANTAITQAVSAIDSQ 212
QY	178	KTNADALNSFIKKVYIONNEQSVFGFTN-----ANVQPSNYSFVAFSDVTP----- 225
DB	213	KTWADTAATTFIKELDSAKLT-PGTTAEQQAAASQPGNYSFVGFSNDVTTGRTGSQEQ 271
QY	226	--VNYKYARTVNGDEPS---SRILANTINSDTDSWIYSLAGTWTKYQFSNYPGPTG 280
DB	272	DLFSWFAKRAKRWTTGLSAGQQLVSSSTPDTDSWIYSLTGAGSKYTLTFTYGPSTA 331
QY	281	YLFPYKLVKAADANNVGLQYKLNNG-NVQOYEFATSTSANNTTANPTPAVDEIKVAKTV 339
DB	332	YLFPYKLVQSSDSNKIGLQYKLNSETLVPITFGNETN---DSGATPAIDDIKVKVT 387
QY	340	LSGLRGQNTIELSVPTGEGNNKVKAPMIGNYILSSNENNAOKI 383
DB	388	LSNLNFGDNTIEFSVATEE---NKVAPMIGNMYLTSSTNNVDKI 428
RESULT	5	
Q9KH13		
ID	Q9KH13	PRELIMINARY; PRT; 584 AA.
AC	Q9KH13;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DE	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE	ADHESIN PMGAL.4 (FRAGMENT).	
GN	PMGAL.4.	
OS	Mycoplasma gallisepticum.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;	
OC	Mycoplasmataceae; Mycoplasma.	
OX	NCBI_TaxID=2096;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=HS;	
RC	Shen Q.C., Bi D.R., Weng C.J.;	
RT	"Sequence analysis of the pmga multigene family of Mycoplasma	
RT	gallisepticum strain HS.";	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
RL	EMBL: AF275312; AAF91415.1; -	
DR	InterPro: IPR001986; EPSP_synbase.	
DR	PROSITE: P500104; EPSP_SYNTHASE.1.	
DR	NON_TER 584	
SQ	SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;	

Query Match	40.3%;	Score 805;	DB 2;	Length 584;	
Best Local Similarity	42.4%;	Pred. No. 3.4e-32;			
Matches	180;	Conservative 72;	Mismatches 113;	Indels 60;	Gaps

QY	10	NPNGQT----	OLEAARWELTDLINAKAMTILASLDYAKIEASLSAYSEATVNNNLNA	65
		:	:	: : : : : :
Db	69	NPNGNTTPEQQAARAKTITD	LLGTEHTNVALYADYAKIQSTLSTAYMTAKTASENTSA	128
		:	:	: : : : : :
QY	66	TLEQKMAKTNLESAINQ	ANTDKTTDFNEHPNLVEAYKALKTTLEQATNLEGLSTAYN	125
		:	:	: : : : : :
Db	129	TLDNLRASATITLQAA	IDKAASNKRTFDSANQPLVTAYNQLKTTLOKTTLSLELSENKYS	188
		:	:	: : : : : :
QY	126	QIRNNVLNKKASLTITK	YLDPLNGTGLDLSNEITANKNINNTLISTINEOKTNADALS	185
		:	:	: : : : : :
Db	189	SIKHLSKLEFDAGSAIA	ARTLDP--TMTGTPVEMASVTKANEDIMTAVSKLTWEKTNADKE-	246
		:	:	: : : : : :
QY	186	NSFTKKYIQNNEOSFV	GTETNAN--VQPSNYSFVAFSAQVTP-----VNYKYART	234
		:	:	: : : : : :
Db	247	NDFEKKP--SKEKLVS	TNDRAHQEQPANWSFAGISVDLTGTGNSQNLPNWNEFAQR	304
		:	:	: : : : : :
QY	235	VWNGD---EPSSRILANT	SITDVSWIYSIAGNTYQFSEFSNYGSPSTGYLFPPYKLVKA	291
		:	: : :	
Db	305	VWTSEGQGTGKTALVS	SPVSATDVSWIYSLAGBGTKVTLSFEYGGPTAFLFYPPYKLVQ	364
		:	: : :	
QY	292	ADANNVGLQYKLNNGV	QVQVEF-----ATST-----S	318
		:	:	: : : : : :
Db	365	ADSSVALQYSLAKTS	SKLINFPAKTPMTNQDSQNGVATTSTEGRSSSEYLVLADEVA	424
		:	:	: : : : : :
QY	319	ANNTTANPTPAVDIEK	IVAKIVLSGLRFGQNTIELSYPTGBGNKNKVPAMIGNIYLSNSN	378
		:	:	: : : : : :
Db	425	AVNNEMPTTVSDINAK	TYLTSGLTFEGENTIEFSVPT-----NKVAPMIGNMYLTSNSG	479
		:	:	: : : : : :
QY	379	NADKI 383		
		:		
Db	480	SQSKI 484		

RESULT	6
Q9KH14	
ID	Q9KH14
AC	PRELIMINARY;
OC	PRT;
DT	01-OCT-2000 (T-EMBLrel. 15, Created)
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE	01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE	ADRESIN PMGAL.3.
GN	PMGAL.3.
OS	Mycoplasma gallisepticum.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC	Mycoplasmataceae; Mycoplasma.
OX	NCBI_TaxID=2096;
EN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-HS;
RA	Shen Q.C., Bi D.R., Weng C.J.;
RT	"Sequence analysis of the PMGA multigene family of Mycoplasma
RT	gallisepticum strain HS.";
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF275312; AAF91414.1; ..
SQ	SEQUENCE 680 AA; 72909 MW; A2ABC650083DFE47 CRC64;

Query Match	38.1%;	Score 761;	DB 2;	Length 680;	
Best Local Similarity	39.8%;	Pred. No. 5.7e-30;			
Matches	164;	Conservative 71;	Mismatches 129;	Indels 48;	Gaps

QY	8	DANPNNGQTOLEARME-----	LTDLINAKAMTILASLDYAKIEASLSAYSEATVNN	60
		:	:	: : : : : :
Db	63	DTNPGGGQGMNATQELV	NNAKALSDLIGESKTVELYADYAKIKADLTSAYAVAKTTS	122
		:	:	: : : : : :
QY	61	NNLNATLEQLKMAKTN	LESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQATNLEGLS	120
		:	:	: : : : : :
Db	123	DSSTTLDQVKATSTLTQ	INTAASDKRFDQNSQKLMAYKVLKDXDTLNKKEAIVMSLN	182
		:	:	: : : : : :

Query Match	38.1%	Score 761;	DB 2;	Length 680;
Best Local Similarity	39.8%	Pred. No. 5.7e-30;		
Matches 164; Conservative	71;	Mismatches 129;	Indels 48;	Gaps 7;

[illegible]

QY 121 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLLDSNEITTTANKNINNTL-----ST 173
Db 183 QKYSAILSEIWAASSTAEIYKQILNPVG-----NLPVVAALNAENKILEAIKEBK 236
QY 174 INEQKTNADALSFIKKVQIONNEOSFVGTFTTANVQPSNYSFVAFSADVTVPVNYKARR 233
Db 237 INSEKSNADLFANYQLYKL--DRTKLMSEGSNNTKQPGNYSFVAYASDIASPNWNFAQR 293
QY 234 TVWNGD-----EPSSRILANTNSITDVSWIYSLAGTNTKYQFSNYSSTGYLFFPKL 288
Db 294 TVWTADSRWTSPPLNQLNSAPLTDVSWIYTLSTGTAKYTLTDFYGPQTYLFFPKL 353
QY 289 VKAADANNVGLQYKLNNGNVQVEF-----ATSTSANNNTANPTPAV 330
Db 354 VKTSD--KVLQYKLNQADVPAIQESEAATAAPAEWTDGQESAETATANEKVNPMPSV 411
QY 331 DEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNNENNAK 382
Db 412 NFINVAKVTLNKLKFSNTIEFVPMQDQNMKNKVPAMIGNIYITSNDENAK 463
RESULT 7
Q49498
ID Q49498 PRELIMINARY; PRT; 702 AA.
AC Q49498;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE PMGAI.3 PROTEIN PRECURSOR.
GN PMGAI.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_taxid=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; Pubmed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
DR EMBL: L28424; AAA62417.1; .
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 25 POTENTIAL.
SEQUENCE 702 AA; 75537 MW; 273E8915FE57B9F CRC64;

Query Match 37.7%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 1.3e-29;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;
QY 10 NPNNGQT-----GLEAARMELTDLINAKAMTLASQDYAKIEASLSAYSEAEVTNNLNA 65
Db 69 NPNNGTPEQQAARKTLTDLTGENTVVALYADYAKIQSTLTAYTAKTASENISA 128
QY 66 TLEQLKAMTNLESAINQANTDKTTFDNEHPNLVZAYKALKTTLEQRTNLEGLSSTAYN 125
Db 129 TLENLRSASTTLQAAIDKAANDKRVFDSVNPQPLVAAYNNLKTTLKSKTSLGLESENKY 188
QY 126 QIRNLVDLYNKASSLITKTLPLNG--GTLLDSNEITTTANKNINNTLS--TINQKTN 181
Db 189 GIKNLSKLFDTGSAITAKTDLPTSGERTLEKVN-----ANNIKMATISPELSKKWKGN 245
QY 182 DALNSFIKKVQIONNEOSFVGTFTTANVQPSNYSFVAFSADVT-----VNYKARRTW 236
Db 246 DAF--NEFEKNPLSKELKSTSTAHNQEPANWSFAKYSVDLTLSNQLPNNWFAQRKW 304
QY 237 NGD--EPSSRILANTN--STIDVSWIYSLAGTNTKYQFSNYSSTGYLFFPKLVAAD 293
Db 305 TSENQOPGKTALVSSPVSATDVSWIYSLAGECTKVTLTFTFYGPDPNAFLYLPKLVKAAD 364

QY 294 ANNVGLQYKLNNGNVQVEF-----ATST-----SAN 320
Db 365 SSSVALQYSLNKTSSKLINFKPAETVSTWTDQSENEVATSTSEARSSYKVLVADAEATS 424
QY 321 NNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNNENNA 380
Db 425 NNMENHTPTVSDINTAKVTLGLTGTENTIEFSVPEG-----KVAPMIGNMYLTSNSEQ 479
QY 381 DRI 383
Db 480 VKI 482
RESULT 8
Q9ZID1
ID Q9ZID1 PRELIMINARY; PRT; 645 AA.
AC Q9ZID1;
DT 01-MAY-1999 (TREMREL. 10, Created)
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE M9 PROTEIN.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_taxid=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; Pubmed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family.";
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF032890; AAC69269.1; .
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;
Query Match 37.3%; Score 746; DB 2; Length 645;
Best Local Similarity 42.3%; Pred. No. 2.9e-29;
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;
QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASQDYAKIEASLSAYSEAEVTN 60
Db 44 DTPNGDQGMMAASQELAAARMGLTVDFDSKAKNLGLYVDYKKTQDTLTAKYDAKATVL 103
QY 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVZAYKALKTTLEQRTNLEGLS 120
Db 104 DNSSSTQNLNEAKRLETAIRTAATSKOTFDEQHAELVKVYEELKTLTSLNETATLAPYA 163
QY 121 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLLDSNEITTTANKNINNTL--STNEOK 178
Db 164 AAQYAGIKHMLSLYDAGKAITTKTLEPVGDP--LTADVVMANTKIVEAIKDEVLPOR 222
QY 179 TNADALSFIKKVQIONNEOSFVGTFTTANVQPSNYSFVAFSADVTVP-----NRYK 230
Db 223 ENATKLADSFVQVLVKEKITGTGVEEAAH--KAOPANTSFVGYSDVTGTGTQTSIPNWDY 281
QY 231 ARRTW--NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSNYSSTGY 281
Db 282 AORTIFTNDEP--RSISNTPADQGTWVQPLSNVSIYSLAGTGAKYTLFTFYGPSTGY 339
QY 282 LYFPYKLVKAADANNVGLQYKLNNGNVQVEFATSTSA-----NNTANPTPAVDEIKVA 336
Db 340 LYFPYKLVNTSDMKLGLEYKLD-----ATEPSAITFGSEQTMNGKPTPVNDINVA 391
QY 337 KIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNNENNAK 383
Db 392 KVTLANLKFGSNKIEFSVPA-----EKVSPMIGNMYLTSNPNWNKI 433
RESULT 9

```
Q9ZHR9          PRELIMINARY;          PRT;          644 AA.
AC Q9ZHR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (MP) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

Query Match          37.3%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 3.3e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

QY 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVN 60
DB 44 DTNPGDGGMMNAASOELAAARMGLTTFVDSKAKNLGLYVDYKKTQDTLTAYDAARTVL 103
QY 61 NNLNATLEOLKMAKTLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAINLEGLS 120
DB 104 DNSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKVYEELATLSNETATLAPYA 163
QY 121 STAYNOIRNLVDLYNKASLLTKTLDPLNGGTLDSNEITANKINNVL--STNEOK 178
DB 164 AAQYAGIKMHLGSLYDAGAKITTKLEPVEGDP-LTASAVMANTRKIVEAIKDEVLPK 222
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NRYK 230
DB 223 ENATKLADFVKQVLVKEKITGVEEAHNS-QPANTSFVGSVDITGTANGQTSIPNNWY 281
QY 231 ARTYV-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSFGYSGTGY 281
DB 282 AQTITNGDEP--RSVNTPDVQGTMAQPLSNVSWIYSLAGTGAKYTLEFYTGSTGY 339
QY 282 LYPYKLVKAADANNVGLQYKLNNGVQVEFATSTSA-----NNTTANPTPAVDIKVA 336
DB 340 LYPYKLVNTSDQVKGLEYKLDN-----ATKPSAITFGSEQTMGKTPTVNDINVA 391
QY 337 KIVSLGRFGQNTIELSVPTGEGNMKNKVPMTGNINYLSSNENNAKDI 383
DB 392 KVTLANLFGSNKIEFSVP-----VEKSPMIGNMVYLSSSPNNWNKI 433

RESULT 10
Q9495          PRELIMINARY;          PRT;          650 AA.
AC Q9495; O08060;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HAEMAGGLUTININ PRECURSOR.
GN PMGAL.1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=S6;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50152.1; -.
KW Signal.
FT SIGNAL          1 25 POTENTIAL.
FT CHAIN           26 650 HAEMAGGLUTININ.
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;
```

```
RC STRAIN=S6;
RA MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50152.1; -.
KW Signal.
FT SIGNAL          1 25 POTENTIAL.
FT CHAIN           26 650 HAEMAGGLUTININ.
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;

Query Match          37.2%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 3.7e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

QY 10 NPNNQG-----TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATVNN 62
DB 51 NPGDGGGMMNAASOELAAARMGLTTFVDSKAKNLGLYVDYKKTQNTLTAYDAARTVL 110
QY 63 LNAITLQKMAKTLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRAINLEGLS 122
DB 111 SSSSTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKVYEELATLSNETATLAPYA 170
QY 123 AYNOIRNLVDLYNKASLLTKTLDPLNGGTLDSNEITANKINNVL--STNEOKTN 180
DB 171 QYAGIKMHLGSLYDAGAKITTKLEPVEGDP-LTAGAVTMANTKIVEAIKDEVLPNPK 229
QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NRYAR 232
DB 230 ATKLAUSFVKQVLVKEKITGVEEAHN-KAQPANYSFVGYSVDITGTANGQTSIPNNWY 288
QY 233 RTYV-NGDEPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSFGYSGTGY 283
DB 289 RTITNGDEP--RSVNTPDVQGTMAQPLSNVSWIYSLAGTGAKYTLEFYTGSTGY 346
QY 284 FPKLVKAADANNVGLQYKLNNGVQVEFATSTANNTTAN-PTPAVDIKVAKIVLSG 342
DB 347 FPKLVNTSDQVKGLEYKLDN---TEPSAITFGNEQTMGKTPTVNDINAKVTLAN 402
QY 343 LRFQNTIELSVPTGEGNMKNKVPMTGNINYLSSNENNAKDI 383
DB 403 LIFGSKNIEFSVPA-----EKVSPMIGNMVYLSSSPNNWNKI 438

RESULT 11
Q9497          PRELIMINARY;          PRT;          649 AA.
AC Q9497;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
```

```

RL  EBBS Lett. 352:347-352(1994).
KW  EMBL; L28424; AAA62416.1; -.
FT  SIGNAL
SQ  SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 37.2%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 4.1e-29;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

QY 8 DAPNNGQ-----TOLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60
   |||||
Db 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKARNLGLYVDYKKTQNTLTTRAYDAAKTVL 107
   :|||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
   :|||
Db 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTDEQHAELVKVYKELKTTLSNETATLAPYA 167
   :|||
QY 121 STAYNQIRNLVDLYNKASSLITKTLPLNGGTLDSNEITTANKNINNL--STINEQK 178
   :|||
Db 168 DAQVAGIKMHLUSGLYDAGKAITTKTLEPEVGDPLTASAVMMANTKIVEAIKDEVLPQK 226
   :|||
QY 179 TN-----ADALSNSIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 226
   :|||
Db 227 ENATKLADSFVKVQVLRKIEKTVGEAHN-KAQPANYSFVGSVDITGTTGTSIPNWDY 285
   :|||
QY 231 ARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGY 281
   :|||
Db 286 AQRITFTNSDEP--RSISNTPADQQTMAQPLSNVSIYSLAGTGAKYTLEFTYGPSTGY 343
   :|||
QY 282 LYFPYKLVKAADANNVGLQKLNNGVQVEFATSTA-----NNTTANPTPAVDEIKVA 336
   :|||
Db 344 LYFPYKLVNTSDQVKLGLEYKLN-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395
   :|||
QY 337 KIVLSGLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
   :|||
Db 396 KYTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 12
Q49468 PRELIMINARY; PRT; 647 AA.
AC Q49468; Q53303;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)
DE HEMAGGLUTININ HOMOLOG PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93162830; PubMed=8432610;
RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmga, a
RT hemagglutinin of Mycoplasma gallisepticum.";
RL Infect. Immun. 61:903-909(1993).
DR EMBL; M83178; AAA02996.1; -.
DR EMBL; S55216; AAB25397.2; -.
FT SIGNAL.
KW SIGNAL.
FT CHAIN
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match 35.7%; Score 713; DB 2; Length 647;
Best Local Similarity 41.8%; Pred. No. 1.2e-27;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

QY 8 DAPNNGQ-----TOLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60

```

```

Db 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKARNLGLYVDYKKTQNTLTTRAYDAAKTVL 107
   :|||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
   :|||
Db 108 DNSSSTQNLNEAKTRLETAIRTAATSKQTDEQHAELVKVYKELKTTLSNETATLAPYA 167
   :|||
QY 121 STAYNQIRNLVDLYNKASSLITKTLPLNGGTLDSNEITTANKNINNL--STINEQK 178
   :|||
Db 168 DAQVAGIKMHLUSGLYDAGKAITTKTLEPEVGDPLTASAVMMANTKIVEAIKDEVLPQK 226
   :|||
QY 179 TN-----ADALSNSIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 226
   :|||
Db 227 ENATKLADSFVKVQVLRKIEKTVGEAHN-KAQPANYSFVGSVDITGTTGTSIPNWDY 285
   :|||
QY 231 ARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNTKYQFSNYPSTGY 281
   :|||
Db 286 AQRITFTNSDEP--RSISNTPADQQTMAQPLSNVSIYSLAGTGAKYTLEFTYGPSTGY 343
   :|||
QY 282 LYFPYKLVKAADANNVGLQKLNNGVQVEFATSTA-----NNTTANPTPAVDEIKVA 336
   :|||
Db 344 LYFPYKLVNTSDQVKLGLEYKLN-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395
   :|||
QY 337 KIVLSGLRFQONTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
   :|||
Db 396 KYTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 13
Q9KH15 PRELIMINARY; PRT; 656 AA.
AC Q9KH15;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-OCT-2000 (TREMREL. 15, Last annotation update)
DE ADHESIN PMGAL.2.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
RT gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91413.1; -.
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;

Query Match 35.2%; Score 704; DB 2; Length 656;
Best Local Similarity 41.0%; Pred. No. 3.4e-27;
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

QY 8 DAPNNGQ-----OLEARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60
   |||||
Db 63 DINPGGONMDSAAQELTAARTALTSILASKANVANVMSDYAKIQNTLIAAYTTAQTQS 122
   :|||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
   :|||
Db 123 QNSATLEQVKNATSAQTALNTANSNKQKQDQHSNLLMSYKNLMATLAKKFAVMTLK 182
   :|||
QY 121 STAYNQIRNLVDLYNKASSLITKTLPLNGGTLDSNEITTANKNINNL--TINEQK 178
   :|||
Db 183 DPYSAILDQINGVSSKGEELVQHTLPVS-GIVPAANTITEETKIEEYISEKTLDQK 241
   :|||
QY 179 TNADALSN--SFI--KKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV----- 226
   :|||
Db 242 NNADQFANYQSTLDKTKLENVEDA-----KKMGQPNYSFVGSVDVTGTSQETTP 295
   :|||
QY 227 NYKVARVTWNGDEPSSRIANTNS-----ITDVSWIYSLAGTNTKYQFSNYPSTGY 278

```



```

Db 296 NWNFAQRAIFTSGNQPKVATTTGGTBDQSTAKPLSDVSWIYSLAGTGAKTTLETTYGGS 355
      | : | | : : : : | : | : | | | | | | | | | | : | : | | |
Qy 279 TGLYFYFKLVKAAADANVGLQYKLN-NGNVQOVEFATSTSANNTTANPTPAVDEIKVAK 337
      | : | | | | | | : | | | | | | : | : | | | | | : | : | | |
Db 356 TGMVLPYFKLVKAND-DVGLQYKLSNETLPIIFGECT---TTNGPAATVENINVAK 409
      | : | | | | | | : | | | | | | : | : | | | | | : | : | | |
Qy 338 IVLSGLRFGONTIELSVPTGEGNNKVAPMIGNIYLSSENENADK 382
      : | : | | | | | | | | : | | | | | | | | : | : | : |
Db 410 VRLTGLAFGKNTIEFSVP-----MSKVAPMIGNMYITSSDTETNK 449
      : | : | | | | | | | | : | | | | | | | | : | : | : |

RESULT 14
O05122 PRELIMINARY; PRT; 703 AA.
AC O05122;
ID O05122;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HAEMAGGLUTININ.
GN PMGAL.9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
RX NCBI_Taxid=2096;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=S6;
RC MEDLINE=95010739; PubMed=7925999;
RX Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum.";
RN FEBS Lett. 352:347-352(1994).
RL [2]
RS SEQUENCE FROM N.A.
RC STRAIN=S6;
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50154.1; -.
DR InterPro; IPR002819; HD.
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73BC5 CRC64;

```

Query Match	34.6%	Score	690.5	DB 2	Length	703			
Best Local Similarity	35.0%	Pred. No.	1.7e-26						
Matches	167	Conservative	75	Mismatches	120	Indels	115	Gaps	12
QY	1	CMSITKKDANP-----	NNGQT-----	QLEARNELDTDLINAK	32				
Db	26	CTSATIPLNTPEPKDPMP	PPSGMGGNGT	PCMDTAQELASAKAAL	171	TNRE	85		
QY	33	AMTLASLDQYAKIEASLS	SEAYETVNNLNLALEQL	KMAKTUNLESAINQ	ANDTKTFD	92			
Db	86	SEKVGLYVDYAKIADLT	AYTVAKTTSDSSTLS	VOVTKATSLQTA	IDAASDQKQFE	145			
QY	93	NEHENLYEAKALKATTLE	QR-ATNLEGLSSPAY	QOIRNLVDLNKASSL	TKTKTDLPLNG	151			
Db	146	QDHKDLMPKSELKTTLS	QKNATVL--LNQPKY	SAILNKINSIIAQGE	EVVIRTDLPVS-	202			
QY	152	GTLLDSNEIITANKNIN	TLS--TINQKTNADALS	NSFIKKVIQNNQSP	QVGTFTNANV	209			
Db	203	GAITPAISITKVNDIEN	KAIASENOLKPKFNAD	AFANYQFKL---DKTK	IMGSTNNMK	259			
QY	210	QPSNYSVFASADYTPV-	-----NYKIARTV	MNGDEPSSRLATNS-	-----	251			
Db	260	QPQNSYFVGYSVGTGH	QSGGTTIPNNFAQRI	YVWSSCAPRASQ	TETQAEPPMSA	319			
QY	252	-----	ITDVSWIYSLA	262					
Db	320	POGVEPAQQQDSDSPK	QASETOEVSPTPA	AEVQAQQADTEQ	PATSQGTPLD	VSWISLS	379		
QY	263	GTNTKYOFSEFSNYG	PTGILYFPYKLVKA	DANNVGLQYKLNN	GVNOVEPAT	TSANNT	322		

```

Db      380 GTDVKYFTTFNYGSPSMAYLYFPYKLVKSD--SVGLQYKLNNNPVALNFGSETNAN-- 435
Qy      323 TANPTPADETKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPAMTIGNIYLVSSNEN 379
Db      436 --GPAASVDNINAKVLANLNFGENTIEFSVP-----MKNKVPAMTIGNIYTSVDVAN 485

RESULT 15
Q9L8D5 PRELIMINARY; PRT; 419 AA.
AC Q9L8D5
DT 01-OCT-2000 (TEmblrel. 15, Created)
DT 01-OCT-2000 (TEmblrel. 15, Last sequence update)
DT 01-OCT-2000 (TEmblrel. 15, Last annotation update)
DE PGKA-LIKE PROTEIN 9.3 (FRAGMENT).
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.F., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel pmgA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum."
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF210770; AAF29525.1; -
FT NCBI 419
FT NT
SQ SEQUENCE 419 AA; 45072 MW; FE5BEB37F2DB3B0C CRG64;

```

[illegible]

Search completed: June 12, 2002, 10:50:40
Job time: 203 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:47:42 ; Search time 17.44 Seconds
(without alignments)
872.522 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSITKDPNNGQTOLEA.....SSNENNADKIPGRRPCTFL 393

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	7.7	682	1 NISP_LACLA	Q07596 lactococcus
2	143	7.2	857	1 AR56-CANAL	P78586 c arg5, 6 pr
3	138.5	6.9	1314	1 SW11-YEAST	P09547 saccharomyc
4	137	6.9	719	1 YM41-YEAST	Q03213 saccharomyc
5	135.5	6.8	1085	1 CUP7-SCHPO	P24339 schizosacch
6	132	6.6	1271	1 Y338-MYCGE	P47580 mycoplasma
7	131.5	6.6	1164	1 BAG-STRA	P27951 streptococ
8	129	6.5	1251	1 RBP2-PLAVB	Q00799 plasmodium
9	128.5	6.4	2334	1 WAPA-SACSU	Q07833 bacillus su
10	128.5	6.4	3712	1 LMA-DROME	Q00174 drosophila
11	128	6.4	995	1 Y109-YEAST	P40442 saccharomyc
12	127	6.4	956	1 YEF3-YEAST	P32618 saccharomyc
13	126.5	6.3	444	1 PST1-YEAST	Q12355 saccharomyc
14	126.5	6.3	1630	1 MSP1-PLAFK	P04932 plasmodium
15	126.5	6.3	1839	1 MSP1-PLAFW	P04933 plasmodium
16	125.5	6.3	1858	1 P3K2-DICDI	P54674 dictyostell
17	124.5	6.2	821	1 L1N5-CAEEL	P45970 caenorhabdi
18	124.5	6.2	1010	1 YK11-CAEEL	P34278 caenorhabdi
19	124.5	6.2	1018	1 FNA8-STAAU	P14738 staphylococ
20	124	6.2	749	1 MAD1-YEAST	P40957 saccharomyc
21	124	6.2	750	1 YK37-YEAST	P42431 saccharomyc
22	124	6.2	823	1 NSP1-SCHJA	P14907 saccharomyc
23	124	6.2	866	1 MYSP-SCHJA	Q05870 schistosoma
24	124	6.2	1159	1 N124-SCHPO	Q09904 schizosacch
25	124	6.2	1433	1 CAT8-YEAST	P39113 saccharomyc
26	123.5	6.2	537	1 ARP-PLAFA	P04931 plasmodium
27	123.5	6.2	1902	1 P2P-LACLC	P15293 lactococcus
28	123	6.2	1260	1 ALS1-CANAL	P46590 candida alb
29	123	6.2	1637	1 MRSP-STAAU	P80544 staphylococ
30	123	6.2	1790	1 US01-YEAST	P25386 saccharomyc
31	122.5	6.1	507	1 FLIC-SALON	Q06974 salmonella
32	122	6.1	1117	1 YN96-YEAST	P53753 saccharomyc
33	122	6.1	1672	1 PMPB-CHLMU	Q9P1Y2 chlamydia m

RESULT 1
NISP_LACLA
ID NISP_LACLA STANDARD; PRT; 682 AA.
AC Q07596;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nisin leader peptide processing serine protease NISP precursor
DE (EC 3.4.21.-)
GN NISP.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIZO R5;
RX MEDLINE=932339683; PubMed=8478324;
RA van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,
Kuipers O.P., de Vos W.M.;
RT "Characterization of the Lactococcus lactis nisin A operon genes
nisp, encoding a subtilisin-like serine protease involved in
precursor processing, and nispR, encoding a regulatory protein
involved in nisin biosynthesis.";
RT J. Bacteriol. 175:2578-2588(1993).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6F3;
RX MEDLINE=94213458; PubMed=8161176;
RA Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,
Hammelmann M., Entian K.-D.;
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
6F3.";
RT Appl. Environ. Microbiol. 60:814-825(1994).
RL [3]
RP 3D-STRUCTURE MODELING
RX MEDLINE=95357326; PubMed=7630881;
RA Siezen R.J., Rollme H.S., Kuipers O.P., de Vos W.M.;
RT "Homology modelling of the Lactococcus lactis leader peptidase Nisp
and its interaction with the precursor of the lantibiotic nisin.";
RL Protein Eng. 8:117-125(1995).
CC -!- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC -!- PATHWAY: LAST STEP OF NISIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
SUBTILASE FAMILY.
CC -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

34 121.5 6.1 425 1 YBVO-YEAST P38272 saccharomyc
35 121.5 6.1 1140 1 YM96-YEAST Q04893 saccharomyc
36 121 6.1 491 1 TOLC-SALEN Q54001 salmonella
37 121 6.1 796 1 YS8A-CAEEL Q09625 caenorhabdi
38 121 6.1 901 1 PIP-LACLA P49022 lactococcus
39 121 6.1 979 1 P115-MYCHR P41508 mycoplasma
40 121 6.1 1111 1 KIF1-YEAST P28742 saccharomyc
41 121 6.1 1902 1 P2P-LACPA Q02470 lactobacill
42 120.5 6.0 776 1 VP4-ROTEH Q02945 equine rota
43 120.5 6.0 800 1 INLA-LISMO P25146 listeria mo
44 120.5 6.0 1565 1 PAC-STRMU P11657 streptococ
45 120 6.0 650 1 PDR4-YEAST P19880 saccharomyc

ALIGNMENTS


```

Query Match      6.98; Score 137; DB 1; Length 719;
Best Local Similarity 19.8%; Pred. No. 1.1;
Matches 105; Conservative 72; Mismatches 175; Indels 178; Gaps 23;

QY  2  MSITKKDANPNNGQTQEAARMETDL-----INAKMTLAS----- 38
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   38  MPTTLKGYRNMNSVNDEAIGINL-DLSLPHISPTTGSBSAGSNATLRNDCNALDGG 96

```

QY 39 -LQDYAKIEA-----SLSSAYSEAEVNNLNNA-----TLEQLKMAKTNLESAL 81
Db 97 LKRTSAISAPCTGPTTTEGKLSNEKRVNSVNSASNSTAGTGRMLUSQSLTN-DSPS 155
QY 82 NOANTDK-----TTFDNEHENLV---EAYKALKTTLEORATNLEGLSSSTAY----- 124
Db 156 NEISTDLQAIQFORMDEMSARMTEEMEEFKNLSKIAEQNTWLVNLKQDNYKYVKNKLNI 215
QY 125 -----NOTRNNL-VDIYNKASSLTKYTDPL--NGTLLDSNEITANKN---- 166
Db 216 KLVAOPSARPSTNNAQKLAIELLNSISAVSAYLQKMNQNGSGRQHTADLCTGDSNTHS 275
QY 167 -----INNTLSTINEOKTNA-----DALNSRFTKKVI 193
Db 276 GINQHTTNGTIDVNTAQLNANQFNSNALNTLPDQOHRNNSVONINOSLPNRQLGPVI 335
QY 194 -----QNNQSFV-GFTFNANQPSNYSFAVSAD-----VTPVNYKYARTVWNGDEPS- 242
Db 336 NQANQNSQOVLIHNTTHQOVNRSPIFPNASTDKPFLNPNKIKRRRTQSNNAST 395
QY 243 -----SRILANTNSITDVSWIYS-----LAGTNKYQF-SFNSYGPSTGY 281
Db 396 NDHASAAQKPIALSPLTNSHNSSTSMNTNSIHSVTSASNSFHDLSLNFQTTAL 455
QY 282 LYFPYKLVKAADANNVGLQKLNNGNVQOVERATSTSANNTTANPTPAYDEIKVAKIVLS 341
Db 456 SLPSALDNASPPPNQNVIPPIINNTQQLSFLSQNLQDSTTS----- 498
QY 342 GLRFQNTIELSVPTGEGNNKVPAMIGNIYLSNNENNADKIPGYRRPOT 391
Db 499 -----EL-LPSGKSGVNT-----NIV---NNRSTLPSYPKPMW 529

RESULT 5
CUT7_SCHPO STANDARD; PRT; 1085 AA.
AC P24339;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Kinesin-like protein cut7.
GN CUT7 OR SPAC25G10.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91015362; PubMed=2145514;
RA Hagan I., Yanagida M.;
RT "Novel potential mitotic motor protein encoded by the fission yeast cut7+ gene.";
RL Nature 347:563-566(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA McLean J., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES; THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO FORM A SHORT SPINDLE THAT ELONGATES TOSPAN THE NUCLEUS AT METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; X57513; CAA40738.1; -.
DR EMBL; Z70691; CAA94636.1; -.
DR PIR; S14032; S14032.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM0129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle; Phosphorylation; Repeat.
FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 436 604 COILED COIL (POTENTIAL).
FT DOMAIN 715 740 COILED COIL (POTENTIAL).
FT DOMAIN 897 955 COILED COIL (POTENTIAL).
FT NP_BIND 159 166 ATP (BY SIMILARITY).
FT REPEAT 987 998
FT REPEAT 999 1010
FT MOD_RES 1011 1011
FT CONFLICT 34 61 PHOSPHORYLATION (BY CDC2) (BY
FT SIMILARITY).
FT SASNPKRREPTIDTGYDRSDTNSPT ->
FT LRAILGNDVSLLLTL (IN REF. 1).
SQ SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;

Query Match 6.8%; Score 135.5; DB 1; Length 1085;
Best Local Similarity 20.9%; Pred. No. 2.2;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;
QY 13 NCQ-TQLEAARMELTDLINAKA-MTLASL-----QDYAKTEASLSSAYSEAEVNNLNAT 66
Db 623 NGYFTLLDNDFASMEELLNTHSNQLLISMTKITEHFQSLDEALQSARSSCAVPNSSLDLI 682
QY 67 LEQLKMAKTNLESAINOANTDTTFDNEHPN-----LVAYKALKTT---LEORATNLES 118
Db 683 VSELKDSKNSLLDALEHSLQDISMSQKLGNGISSLELLEQDKMESYQLVQELRSLYN 742
QY 119 LSTAYNQIRNNLVLYNKASSLTKYTDPLNGTLLDSNEIT-----TA 163
Db 743 LQHTHEESQKELMYGVRNDIDALVKTTCTSLNDADILSDIYSDQSKPESKQODLIANI 802
QY 164 NNNINTLSTINEQ-KTNADAL-----SNSPIKK---VIONNEQSFVCTFTNANVQPSN 213
Db 803 GKIVSNFLQEQNESLYTKADILHSHLNDTNSNIRKANEIMNRRSEFL---RNA----- 853
QY 214 YSFVAFSADVTVPVNYKYARTVWNG-----DEPSSRLTANTNSTIDVSWIYSLA----- 262
Db 854 -----ASQAEIVGANKERIOKTWENGSLDLSKSKATHSNRSMYD-----HCLALAESQK 905
QY 263 GNTKYQ-----FSFNYGPTGYLYFPYKLVKAADANNVGL--QYKLNNGNYQ 309
Db 906 GVNLEVTDLRLQLQKVEHSEDNTEKHQOL---LDLLESVLGNNNDLDSIKTPHTLEQ 962
QY 310 QVE-----FATSTSANNT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 350
Db 963 KITDHLVKGTTSILANTNELLGLGDESCLNLETTIEDTSLVKLETTGDTPTPSKRELPA 1022
QY 351 -----ELSVPTGEGNNKVPAMIGNIYLSNNENNADKIPGYREP 389
Db 1023 WTRDSSLIKETNLDLSDKKFVRETTSSNQTNPEPDV---YDKP 1064
RESULT 6
Y338_MYCGE
ID Y338_MYCGE STANDARD; PRT; 1271 AA.
AC P47580;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MG338 precursor.

Qy	251	SIT--DVSWIYSLAGTWTKYQFSNYGSPSTGYLYPKFKLVKAADANNVGLOKLNNGV	308
Dq	533	NFSDKNVSESVOLFSGNSFR-SWANRN-TTLKIYTALTTMLENGTSEN-----NNGQR	593
Qy	309	QQVEFATSTSANNTANPT 327	
Dd	584	DVCDLARKLLKNNTNLSET 602	
RESULT 7			
BAG_STRAG	ID	BAG_STRAG STANDARD; PRT; 1164 AA.	
AC	P27951;		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Iga FC receptor precursor (Beta antigen) (B antigen).		
GN	BAG.		
OS	Streptococcus agalactiae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1311;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 38-48.		
RC	STRAIN-LA239;		
RX	MEDLINE=91312121; PubMed=1857207;		
RA	Jerlstrom P.G., Chhatwal G.S., Timmis K.N.;		
FT	"The Iga-binding beta antigen of the c protein complex of Group B streptococcus": sequence determination of its gene and detection of two binding regions."		
RL	Mol. Microbiol. 5:843-849(1991). [2]		
RP	IDENTIFICATION OF IG-LIKE DOMAIN.		
RX	MEDLINE=97035265; Pubmed=8880921;		
RA	Bateman A., Eddy S.R., Clothia C.;		
FT	"Members of the immunoglobulin superfamily in bacteria.";		
RL	Protein sci. 5:1939-1942(1996).		
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.		
CC	-I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.		
CC	-I- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS IN THE REGION OF THE MEMBRANE ANCHOR.		
CC	This SWISS-PROT entry is copyright. It is produced through a collabor between the Swiss Institute of Bioinformatics and the EMBL Outstat the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for comm entities requires a license agreement (See http://www.isb-sib.ch/anno or send an email to licens@isb-sib.ch).		
DR	EMBL; X59771; CAA42442.1; -		
DR	PIR; S15330; FCSOAG.		
DR	InterPro; IPR001899; Gram_pos_anchor.		
DR	InterPro; IPR003599; Ig.		
DR	Fam; PF00746; Gram_pos_anchor; 1.		
DR	SMART; SM00409; IG; 1.		
KW	PROSITE; PS00343; GRAM_POS_ANCHORING; 1.		
KW	Cell wall; Transmembrane; Receptor; Repeat; Signal; Immunoglobulin domain.		
FT	SIGNAL 1 37		
FT	CHAIN 38 1164		
FT	DOMAIN 38 1131		
FT	EXTRACELLULAR (POTENTIAL);		
FT	MEMBRANE ANCHOR (POTENTIAL).		
FT	CYTOPLASMIC (POTENTIAL).		
FT	IG-LIKE DOMAIN.		
FT	IGA-BINDING (POTENTIAL).		
FT	IGA-BINDING (POTENTIAL).		
FT	PRO-RICH REPEATS.		
FT	CONSERVED IN GRAM-POSITIVE COCCI SURFACE PROTEINS.		
SQ	SEQUENCE 1164 AA; 131051 MW; 65DE94AEF720A547A CRC64;		

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
CC IN CELL PATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING
CC STRUCTURE.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY
CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
CC -1- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
CC DEVELOPMENT AT 10-12 HOURS.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV)' IS NOT
CC SIMILAR TO LAMININ DOMAIN IV).
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M96388; AAA28662.1; -
CC EMBL; L07288; AAC37178.1; -
CC EMBL; M75882; AAA28661.1; -
CC HSSP; P02468; 1TLE
CC Flybase; FBgn0002526; LANA.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001886; LamNT.
CC InterPro; IPR000034; Laminin_B.
CC InterPro; IPR002049; Laminin_EGF.
CC InterPro; IPR001791; Laminin_G.
CC InterPro; IPR001230; Prenyltn.
CC Pfam; PF00052; laminin_B; 1.
CC Pfam; PF00053; laminin_EGF; 20.
CC Pfam; PF00054; laminin_G; 5.
CC Pfam; PF00055; laminin_Nterm; 1.
CC PRINTS; PR00011; EGF_LAMININ.
CC ProDom; PD002082; LamNT; 1.
CC ProDom; PD003031; Laminin_B; 1.
CC SMART; SM00180; EGF_Lam; 17.
CC SMART; SM00001; EGF_Like; 1.
CC SMART; SM00281; LamB; 1.
CC SMART; SM00282; LamG; 5.
CC SMART; SM00136; LamNT; 1.
CC PROSITE; PS00022; EGF_1; 17.
CC PROSITE; PS01186; EGF_2; 5.
CC PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
CC PROSITE; PS00025; LAM_G_DOMAIN; 5.
CC Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
CC SIGNAL 1 22 POTENTIAL.
CC CHAIN 23 3712 LAMININ ALPHA CHAIN.
CC DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).
CC DOMAIN 273 332 LAMININ EGF-LIKE 1.
CC DOMAIN 333 402 LAMININ EGF-LIKE 2.
CC DOMAIN 403 447 LAMININ EGF-LIKE 3.
CC DOMAIN 448 494 LAMININ EGF-LIKE 4.
CC DOMAIN 495 540 LAMININ EGF-LIKE 5.
CC DOMAIN 541 586 LAMININ EGF-LIKE 6.

FT DOMAIN 587 631 LAMININ EGF-LIKE 7.
FT DOMAIN 632 676 LAMININ EGF-LIKE 8.
FT DOMAIN 677 731 LAMININ EGF-LIKE 9.
FT DOMAIN 732 784 LAMININ EGF-LIKE 10.
FT DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 816 1374 DOMAIN IV'.
FT DOMAIN 1375 1420 LAMININ EGF-LIKE 12.
FT DOMAIN 1421 1465 LAMININ EGF-LIKE 13.
FT DOMAIN 1466 1513 LAMININ EGF-LIKE 14.
FT DOMAIN 1514 1564 LAMININ EGF-LIKE 15.
FT DOMAIN 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).
FT DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).
FT DOMAIN 1809 1858 LAMININ EGF-LIKE 17.
FT DOMAIN 1859 1916 LAMININ EGF-LIKE 18.
FT DOMAIN 1917 1969 LAMININ EGF-LIKE 19.
FT DOMAIN 1970 2016 LAMININ EGF-LIKE 20.
FT DOMAIN 2017 2063 LAMININ EGF-LIKE 21.
FT DOMAIN 2064 2111 LAMININ EGF-LIKE 22.
FT DOMAIN 2112 2671 DOMAIN II AND I.
FT DOMAIN 2672 2868 LAMININ G-LIKE 1.
FT DOMAIN 2876 3048 LAMININ G-LIKE 2.
FT DOMAIN 3055 3223 LAMININ G-LIKE 3.
FT DOMAIN 3349 3528 LAMININ G-LIKE 4.
FT DOMAIN 3534 3709 LAMININ G-LIKE 5.
FT DOMAIN 2178 2249 COILED COIL (POTENTIAL).
FT DOMAIN 2301 2321 COILED COIL (POTENTIAL).
FT DOMAIN 2376 2450 COILED COIL (POTENTIAL).
FT DOMAIN 2541 2676 COILED COIL (POTENTIAL).
FT DOMAIN 3270 3296 POLY-THR.
FT DISULFID 273 282 BY SIMILARITY.
FT DISULFID 275 296 BY SIMILARITY.
FT DISULFID 298 307 BY SIMILARITY.
FT DISULFID 310 330 BY SIMILARITY.
FT DISULFID 333 342 BY SIMILARITY.
FT DISULFID 335 367 BY SIMILARITY.
FT DISULFID 370 379 BY SIMILARITY.
FT DISULFID 382 400 BY SIMILARITY.
FT DISULFID 403 414 BY SIMILARITY.
FT DISULFID 405 421 BY SIMILARITY.
FT DISULFID 423 432 BY SIMILARITY.
FT DISULFID 435 445 BY SIMILARITY.
FT DISULFID 448 460 BY SIMILARITY.
FT DISULFID 450 468 BY SIMILARITY.
FT DISULFID 470 479 BY SIMILARITY.
FT DISULFID 482 492 BY SIMILARITY.
FT DISULFID 495 507 BY SIMILARITY.
FT DISULFID 497 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 528 538 BY SIMILARITY.
FT DISULFID 541 553 BY SIMILARITY.
FT DISULFID 543 560 BY SIMILARITY.
FT DISULFID 562 571 BY SIMILARITY.
FT DISULFID 574 584 BY SIMILARITY.
FT DISULFID 587 599 BY SIMILARITY.
FT DISULFID 589 605 BY SIMILARITY.
FT DISULFID 607 616 BY SIMILARITY.
FT DISULFID 619 629 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT DISULFID 634 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 677 691 BY SIMILARITY.
FT DISULFID 679 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 714 729 BY SIMILARITY.
FT DISULFID 732 746 BY SIMILARITY.
FT DISULFID 734 753 BY SIMILARITY.
FT DISULFID 755 764 BY SIMILARITY.
FT DISULFID 767 782 BY SIMILARITY.
FT DISULFID 1375 1387 BY SIMILARITY.
FT DISULFID 1377 1394 BY SIMILARITY.
FT DISULFID 1396 1405 BY SIMILARITY.

Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).

EMBL; Z46921; CAA87023.1; --
SGD; S0001431; YII169C.
InterPro; IPR004089; Chemotaxis_transducer.
InterPro; IPR000727; T_SNARE.
Hypothetical protein; Signal.

	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	995	HYPOTHETICAL PROTEIN YII169C.
FT	CARBOHYD	28	28	N-LINKED (GLCNAC..)
FT	CARBOHYD	35	35	N-LINKED (GLCNAC..)
FT	CARBOHYD	468	468	N-LINKED (GLCNAC..)
FT	CARBOHYD	664	664	N-LINKED (GLCNAC..)
SQ	SEQUENCE	995 AA;	99735 MW; F63E287A03F137EC CRC64;	(POTENTIAL).

Query Match	6.4%	Score 128;	DB 1;	Length 995;
Best Local Similarity	16.1%;	Pred. No. 5;		
Matches	73;	Conservative 100;	Mismatches 152;	Indels 128; Gaps 15;

QY	3	SITKDPNPNGQTGLEARNMELTDLNAKAMTWLASIQDYAKIASISLSAYSEATVNNN	62
Dd	136	SVSOGSSASDVSSVSQSASSADVVSSVSQSASSAD-----VSSVSQSASTSD	188
QY	63	LNAYLEQLKWAKTNLESAIQAANTDKTFDEHNPLEVEYKALKATTLEORATNLEGSLST	122
Dd	189	VSSVSQSASSASDVSSVSQSASSADVSSVSQSASSTSDDVSSVSQSASSTSGVSSS	248
QY	123	AYNQIRNLVDLYNKASSLIKTLDPLNGGTTLLDSNEITTANKNIINTLTINPKTNAD	182
Dd	249	GSOVS--ASGSSESFPPOST-----SSASTAGSATSNLSLSITSASSAS	293
QY	183	ALSNSFKKK-----VIQNQQSF---VG	202
Dd	294	ATASNLSSSDGTYLPPTTTISGDLTLTGKVIAEGVVAAGAKLTLDGDKYFSADLK	353
QY	203	TFTNAVQPNSYSPFAVSADVPVNYKYARTTVNGDE---PSRILANTNTSIDWSIW	259
Dd	354	YIGDLLVKKEITYPGTHFDISGENFDVTGN-FNAESAATSASIYSFTSPSSFDNSGDI	411
QY	260	SLA-GTKTYQFSFNTPGPSTGYLFPPYKLVAADANNVLQYLKNNGWQQOVFAST-	317
Dd	412	SLSLSKKGKEVFPSYNSGAFSF-----SNAI----LNGGSVGLQRDDTE	456
QY	318	-SANNTANPTPAVEIKVAKIVLSGRFGONTIELSVPTGEHMNVKAPMIGNIYL---	373
Dd	457	GSVNNGEIN-----LDSGSTYYIVEPVSGKGTWNIIIS---GNLYLHP	496
QY	374	-----SSNENNADKIP--GY	386
Dd	497	DTFTGQTVWFKGGVLAVIDPETNATPIPVVGY	529

RESULT 12				
YEF3_YEAST	ID	YEF3_YEAST	STANDARD;	PRT; 956 AA.
AC	P32618;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Hypothetical 106.1 kDa protein in GLY1-GDAL intergenic region.			
OS	YELO43W OR SYGP-ORF14			
OC	Saccharomyces cerevisiae (Baker's yeast).			
CC	Eukaryota; Fungi; Ascomycota; Saccharomycetina; Saccharomycetes;			
CC	Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae;			

OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Allen E., Araujo R.,
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S28C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norrgren R., Oefner P., Oh C.,
 RA Petel F.A., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FIBRONECTIN TYPE III DOMAIN
 RX MEDLINE=97148176; PubMed=8994808;
 RA Bateman A., Chothia C.;
 RT "Fibronectin type III domains in yeast detected by a hidden Markov
 model.";
 RL Curr. Biol. 6:1544-1546(1996).
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U18779; AAB64999.1; -
 DR PIR; S30834; S30834.
 DR SGD; S0000769; YEL043W.
 DR InterPro; IPR003961; FN_III.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 KW Hypothetical protein.
 FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
 SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;
 Query Match 6.4%; Score 127; DB 1; Length 956;
 Best Local Similarity 19.7%; Pred. No. 5.4;
 Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;
 QY 4 ITRKDAMPNGQTOLEARMELTDL-----INA-KAMTLASLDQYAKIEASLSAYSEA 56
 DB 337 LNTESNVAVINKEISLQNEISKEESNRRLNASKSLTISIVVANVENDKPIASGEL 396
 QY 57 ETVNNLN-ATLE-----OLKMAKTNLGSAINQANTDKTTFDNEHPLNVEAYKALK 106
 DB 397 SAVLKLNDFTLEKNGFLSNAGEEFLSKLADSSLIKIKQELSIDQE---LEANWKLQR 453
 QY 107 TTLEQRATNLEGLSSATYQINRNVDLYNKASLITKTL-DPL--NGVLLDSNEITTA 163
 DB 454 SNLLKKISALE-----NQF--NEMSLNR--NLKTKLMVQPKNGKNGDSLAATNSNNA 502
 QY 164 NKN-----INTLSTINEQKTNADAL----- 184
 DB 503 EKNRSSGSIQLPLNNMSRTGSDILSNNNKSN--NSNADSAAPPLRLHNPVSYSPSNEP 560
 QY 185 ---SNSPIKKVIONNE-QSFVGTFTNAN---VQPSNYSFVA-----FSADVTPVNYK 229
 DB 561 IQPSSSLSLQTDQTDNRNMLSNHSSNNENKQPSYSHALPTTANATATATATN-G 619
 QY 230 YARTVWNG---DEPSSRIIA-----NTNSIFDVSWIYSLAGTNTKYQFSFNYG 276
 DB 620 HSRSNLWTTAQAPQSHQVSTELDQAFEDNANHLL--ISGLQNMIVDETDYDPNINISYS 677

QY 277 -----PSTGYLYFPKLYKRAADANNVGVLOYKLNNGVQ 309
 DB 678 KGFTDDELNDYWTQKQPVRSSTNESLFTTGTTPMSSYK-----ANPVISY---SSSHLR 729
 QY 310 QVEFATSTSNANTTANPTPAVDEKVKIVLSGLRF-----QNTIELS-- 353
 DB 730 QTSNATNTNPMIPQSLAATLNDPSLQSFVRSVSGFYSAQPANSLQNNINGNETENISPR 789
 QY 354 -----VPTGEGNMKNVPMICNIGLYLSNENNADKI 383
 DB 790 ISSDFNLVLPNLSPLNSNDVPVPGNVTTLTPSHSNILTMNHQPTADNI 838
 RESULT 13
 PST1_YEAST
 ID PST1_YEAST STANDARD; PRT; 444 AA.
 AC Q12355;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Protoplast secreted protein 1 precursor.
 GN PST1 OR YDR055W OR D4214 OR YD9609.09.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96381250; PubMed=8789263;
 RX Brandt P., Ramlow S., Otto B., Bloecker H.;
 RA "Nucleotide sequence analysis of a 32,500 bp region of the right arm
 RT of Saccharomyces cerevisiae chromosome IV.";
 RL Yeast 12:85-90(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S28C / AB972;
 RC Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
 RA Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN IDENTIFICATION.
 RP MEDLINE=99251092; PubMed=10234784;
 RX Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
 RA "Two-dimensional analysis of proteins secreted by Saccharomyces
 RT cerevisiae regenerating protoplasts: a novel approach to study the
 RT cell wall.";
 RL Yeast 15:459-472(1999).
 RN [4]
 RP GPI-ANCHOR.
 RX MEDLINE=20469049; PubMed=11016834;
 RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
 RT "Up-regulation of genes encoding glycosylphosphatidylinositol
 RT (GPI)-attached proteins in response to cell wall damage caused by
 RL disruption of PKS1 in Saccharomyces cerevisiae.";
 RL Mol. Gen. Genet. 264:64-74(2000).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
 CC SECRETED BY REGENERATING PROTOPLASTS.
 CC -1- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X84162; CAA58971.1; -
 DR EMBL; Z74351; CAA98873.1; -
 DR EMBL; Z49209; CAA98084.1; -
 DR SGD; S0002462; PST1.
 DR COMPLEYEST-2DPAGE; Q12355; -
 KW Glycoprotein; Membrane; GPI-anchor; Signal.
 FT SIGNAL 1 19 POTENTIAL.

```

FT CHAIN 20 ? PROTOPLAST SECRETED PROTEIN 1.
FT PROPEP ? 444 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 356 416 SER-RICH.
FT CARBOHYD 57 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 444 AA; 45776 MW; 230F60CACAS5921A4 CRC64;

Query Match 6.3%; Score 126.5; DB 1; Length 444;
Best Local Similarity 18.1%; Pred. No. 2.3;
Matches 73; Conservative 79; Mismatches 146; Indels 105; Gaps 15;

QY 16 TOLEARME-LTDLINAKMTLASQDYAKIEASLSAYSEAEVTNNLNATLEQKMAK 74
Db 88 TSFAADSLSEITSLQSLTILT-----SASFGSLQSDYSIKLITLPAISSFT 136
QY 75 TNLESAINOANDKTFEDNEHNLVEAYKALKTTLEQRATNLEGLSS--TAYNQIRNNLV 132
Db 137 SNIKSANNIYISDTSL-----QSDGFSALAKVWFVNNNKKLTSIKSPVETVSDSL 190
QY 133 DLYNKASSLITKTLDP---NGGTLDSNEITANKN-----INNLTSTINEQKTN 181
Db 191 FPFNGNQTKI--TFDDLVWANNISLTDVHSVFANLQKINSLSLGFNNSSISLNFETKLT 248
QY 182 -----DALNSFIKK-----VIQNN-----EOSFVGTFIN 206
Db 249 IGQTFISVNDYKLLSFLNSLTIGGALVANNLTGLQKIGGLDNTTIGGTLVEVGNFTS 308
QY 207 ANVQPSNYSFAVSFAFDPVNTKYARRTVWNGDEPSSRLANTNSITDVSWIYSLAGTWT 266
Db 309 LNL--DSLKSVKGADVESKSNFSCNAL-----KALQKKGKIGESFVCKNGASST 358
QY 267 KYQFSFSNYGPTGYLYFPYKLVKAADANNVGLQKLNNGNQVQVEFAFSTSANNTAMP 326
Db 359 SVKLSSTSKSQSS-----QTTAKVSKSSKAEKKFTSGDIKAAASASSVSSSGASS-- 410
QY 327 TPVDEIKVAKIVLSGLRPGQNTIELSVPTGEGNNKVPMTG 369
Db 411 -----SSSKSKGNAAIMAPIGO-----TTPLVG 434

```

```

RESULT 14
MSPL_PLAFK
ID MSPL_PLAFK STANDARD; PRT: 1630 AA.
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman N., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
RT "Polymorphism of the precursor for the major surface antigens of

```

```

FT Plasmidium falciparum merozoites: studies at the genetic level.*;
RL EMBO J. 4:3823-3829(1985).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RA Pan W., Tolle R., Bujard H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03371; CAA27070.1; -.
DR PIR; A25120; SAZQK1.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

```

```

Query Match 6.3%; Score 126.5; DB 1; Length 1630;
Best Local Similarity 20.7%; Pred. No. 11;
Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

QY 36 LASLDQYAK-----TEASLSAYSEAEVTNNLNATLEQKMAKTNLESAINQANDKRT 90
Db 226 VGKMDYIKKKRTIENINELIEESKKTIDKKNKATKEEK-----KKLYQAQYDLSI 278
QY 91 FONEHNLVEAYKALKTTLEQRATNLEGLSSSTAYNOIRNNLVLYNKASSLITKTLDP 150
Db 279 YNKQ---LEEAHN-LISVLEKRTDLUK-----NENIKELDKINEI--KNPPAN 323
QY 151 GG-----TLDSN-EITFANKNNNTLSTINEQKTNADAL-----SNSFIK 190
Db 324 SGNTPTLLDKKKKEEKEKEIAKTI---KFNIDSFLTDPLELEYLYREKNNKIDIS 380
QY 191 KVIQNNQSFVGTFTNANVQPSNYSFV-----AFSADVTVPNTKYARRTVWNGDEPS 242
Db 381 AKVETKESTPEPNEYPNGVTYPLSYNDINNALNELNSFGDLINFPDYT-----KEPS 431
QY 243 SRLAN-----TNSITDVSWIYSLAGTNTYKQFS 271
Db 432 KNIYTDNERKKTINEIKEKIEKKIESDKSYEDRSKSLANDITKEYEKL-LNEIYDSK 490
QY 272 FSNYIGSTGY-----LYFPYKLVKAADANNVGLQKLNNGNQVQVEFAFSTSANNTAMP 326
Db 491 FNNIDLTNFERKMMGKRYSYKVEKLTH--HTPFASYENSKHNLEKLTALKYMEDYSLRN- 548

```

QY 327 TPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMNVAPMIGNIYLSNENNADK 382
 Db 549 IVVEKELKYKNIISKI---ENEIETLVE---NIKDEQLFEKKTIDKDNKPDE 597

RESULT 15
 MSPL_PLAFW STANDARD; PRT; 1639 AA.
 ID MSPL_PLAFW
 AC P04933;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMMA) (P195).
 GN MSP-1.

OS Plasmodium falciparum (isolate Wellcome).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5848;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86014335; PubMed=2995820;
 RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
 RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
 RA Freeman R.R.;
 RT "Primary structure of the precursor to the three major surface
 RT antigens of Plasmodium falciparum merozoites.";
 RL Nature 317:270-273(1985).
 RN [2]

RP REVISIONS.

RA Holder A.A.;

RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 (Potential).

CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; X02919; CAA26676.1; -

DR FIR; A24594; A24594.

DR InterPro; IPR000561; EGF-like.

DR Pfam; PF00008; EGF; 1.

KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;

KW Transmembrane; GPI-anchor.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.

FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).

FT SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match

Best Local Similarity 6.3%; Score 126.5; DB 1; Length 1639;

Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

QY 36 LASLDQYAK-----IEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQANTDKTT 90
 Db 235 VGMEDYIKKNKKTIENTINELIEESKKTIDKNKATKEEK-----KKLYQAYDLSI 287
 QY 91 FDNEHPNLVAYKALKTTLEQRATNLEGLSTAYNQIRNNLVLDLYNKASSLIITKLDPLN 150
 Db 288 YNKQ---LEEAHN-LISVLEKRIDTLKK-----NENIKELDKINEI--KNPPAN 332
 QY 151 GG-----TLDSN-EITTANKNNINNTLSTINEQKTNADAL-----SNSFIK 190
 Db 333 SGYTPNVLDDKNKIEHEKEIEIAKTI---KFNIDSLFTDPLELEYLYREKKNKIDIS 389
 QY 191 KVIONNEQSVGFTTANVQPSNYSFV-----AFSADVTVPVNYKYARTVWNGDEPS 242
 Db 390 AKVETKESTPEPNEYPNGVTYPLSYNDINNALNELNSFGDLINPDYT-----KEPS 440
 QY 243 SRLAN-----TNSITDVSWIYSLAGTNKYOFS 271
 Db 441 KNIYTDNERKKFINEIKEIKIEKKIESDKSYEDRSKSLNDITKEYKL-LNEIYDSK 499
 QY 272 FSNYGPSTGY-----LYFPYKLVKAADANNVGLQYKLNNNGVQOQVEFATSTSANNTTANP 326
 Db 500 FNNIDLTNFEKMMGKRYSKVEKLTH-HNTFASYSKHNLEKLTALKYMEDYSLRN- 557
 QY 327 TPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMNVAPMIGNIYLSNENNADK 382
 Db 558 IVVEKELKYKNIISKI---ENEIETLVE---NIKDEQLFEKKTIDKDNKPDE 606

Search completed: June 12, 2002, 10:51:05
 Job time: 203 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:46:22 ; Search time 29.99 Seconds
(without alignments)
1259.190 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456
Perfect score: 1998
Sequence: 1 CMSTTKDANPNNGQTLEA.....SSNENNADKIPGYPGCTFL 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	41.6	702	2 S48754	major surface prot
2	754	37.7	702	2 S48753	major surface prot
3	744	37.2	650	2 S48751	major surface prot
4	743	37.2	649	2 S48752	major surface prot
5	713	35.7	647	2 A49218	hemagglutinin homo
6	575.5	28.8	386	2 S48755	major surface prot
7	189.5	9.5	320	2 S51560	major surface prot
8	175	8.8	6713	2 B89921	hypothetical prote
9	173	8.7	1302	1 JG6009	surface-located me
10	165.5	8.3	661	2 AG2422	hypothetical prote
11	164.5	8.2	2481	2 D90011	FmtB protein [impo
12	163.5	8.2	3890	2 C89921	hypothetical prote
13	159	8.0	933	2 S41539	fibrinogen-binding
14	156	7.8	807	2 B71605	hypothetical prote
15	155	7.8	1072	2 A86827	hypothetical prote
16	155	7.8	4688	2 F82885	hypothetical prote
17	154.5	7.7	682	2 S44131	subtilisin-like pr
18	154.5	7.7	1365	2 T30822	lmpI protein - Myc
19	151.5	7.6	1237	2 D71850	probable outer mem
20	151	7.6	624	2 PC6003	surface membrane p
21	150.5	7.5	1051	2 T18351	lmpI protein - Myc
22	145	7.3	2399	2 H71879	toxin-like outer m
23	144	7.2	1107	2 AC0976	probable autotrans
24	143.5	7.2	751	2 T40462	ser-lys rich hypot
25	142	7.1	135	2 B49218	hemagglutinin homo
26	142	7.1	2660	2 E85822	probable invasiv z
27	141.5	7.1	1645	2 F96907	phage-related prot
28	140.5	7.0	3194	2 D71917	toxin-like outer m
29	139.5	7.0	1487	2 AG2560	hypothetical prote

30	139.5	7.0	2401	2 T28676	rhostry protein -
31	138.5	6.9	1314	1 TNBYR6	transcription regu
32	137.5	6.9	820	2 T17519	cell surface antig
33	137.5	6.9	926	2 AE1130	conserved hypothet
34	137.5	6.9	989	2 D89852	fibrinogen-binding
35	137.5	6.9	1524	2 S68553	surface layer prot
36	137.5	6.9	4152	2 T31102	filamentous hemag
37	137	6.9	719	2 S55119	hypothetical prote
38	136.5	6.8	568	2 E97066	membrane associate
39	136.5	6.8	1073	2 S14032	kinesin-related pr
40	136.5	6.8	2271	2 F90073	hypothetical prote
41	136	6.8	5005	2 F82884	hypothetical prote
42	135.5	6.8	1085	2 T38378	kinesin-like prote
43	135.5	6.8	1238	2 A64596	hypothetical prote
44	135	6.8	2269	2 T28677	rhostry protein -
45	134.5	6.7	589	2 B97806	hypothetical prote

ALIGNMENTS

RESULT 1

S48754

major surface protein (clone pmGAL.4) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48754
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollock, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S48754
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-702 <MAR>
A:Cross-references: EMBL:L28424
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 41.6%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 3.2e-38;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY	8	DANPNNG-----QTQLEAARMELTDLINAKAMTSLAQDYAKTEASLSAYSAETVNN	61
Db	68	NTNPGNGGGTDNAAQQLAAAKKELSDLLATQNSLSTYADYANIQNTLTAAATKSTSD	127
QY	62	NLNATLEQLKMAKYNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS	121
Db	128	NTSATLEQVKASATSTLTQTAIDTAASSKTSFDEKNPELIKAYNALKETLKKWRNSLSGLTD	187
QY	122	TAYNQRNNVLVDLYNKASSLITTKTLDPLNGTLLDSNEITTANKNNINNTLTSTNEQTNA	181
Db	188	SNFATKTNLTALYQSGKDIIVTKTLDPLM-GTAINLSAVSQATNISNAVSKUETWTNA	246
QY	182	DALNSNFIKKVIQNNESFVGTETNANVPQSNYSFVAFSADVTP-----VNTKYARVT	235
Db	247	TVLATSFVKVLYKKNLTGIDT-TNNOEQPGNYSFVGYSDVTGSDNARNPNSFARKV	305
QY	236	WNGD-----EPSSRIILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLFPYKLVK	290
Db	306	WTSNTDILSQPAEGENQOSAPDVSWIYNTLGMGAKYSLTFTNYGSPSTGLYFPYKLVN	365
QY	291	AADANNVGLQYKLNNGNVQVEFATS-----TSANN-----TTANPTPAVD	331
Db	366	SSSDSKVALEYKLNESAVKTIIDFSPQTSFVADSTARNRSTAAPQAQSGTEINPAPTLD	425
QY	332	EIKVAKTVLSGLRFGONTIELSVPT-TGEGNNKVPAMIGNIYLSNENNAADK	382
Db	426	DIKIAKVTLSNLFKSGSTIEFVSPTTAKEGTSKVAPMIGNMYLTSSDRDYNK	477

A; Residues: 1-320 <MAR>
A; Cross-references: EMBL:L28424; NID:g535687; PID:AA62415.1; PID:g535688
C; Genetics:
A; Genetic code: SGC3

	Query Match	9.5%	Score 189.5	DB 2	Length 320
	Best Local Similarity	48.9%	Pred. No. 0.0015		
	Matches	46	Conservative 11	Mismatches 26	Indels 11
				Gaps	
QY	291	AADANNVGLQKLNNGNVQVEFTST	SANNNTANPTPA	VDDEIKVAKIVLSGLRFGQNTI	350
		: : : : :			
Db	41	SADSN-----PTNQENSQSQAAPESA	-----MNETPTVDGINVAKVTLTDLAFGSNTI	90	
QY	351	ELSPV-TGEGNNKVAPMIGNIYLSSNENNADKI	383		
		: : : :			
Db	91	ELSPVTTDEEGTSKVPAMIGNIYITSDNQORKI	124		

RESULT 8

B89921
hypoetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
R:Accession: B89921
C:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Savano, T.; Inoue, R.; Kafo, C.; Sekimizu, K.;
C:Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701232; PIDN:BAB42527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Query Match	8.88;	Score 175;	DB 2;	Length 6713;
Best Local Similarity	22.63;	Pred. No. 0.54;		
Matches 100;	Conservative	66;	Mismatches 179;	Indels 98;
				Gaps 17;

Qy	8	DANPNNGQTQLEARMELTDLINAKAMTILASLDQYAKIEASISLAYSEAEET-----VNNN	62
		: : : : : : : : : :	
Db	492	EAVSNTG-TQLNTAMANLQNGINDKANTLAS-ENTHDADSOKKTYTQAVTNAENILNKN	549

QY 123 AYNQIRNLVDLYNKAS-SLITKTLDPLNG--GTLLDSNEITTANKNINNTTSTINEOKT 179
::: : | | : : ||| ||| : : | | | |
Db 600 QKDCLKQOVQRQNAVGVDTVKSSANTLANGAMGTLSENSIDONTATKRGONYI.DATERNKT 659

QY 180 -----NADALNSFTKKVIONNEQSFGTETTNANVPSPSYSFVAPS 220
| | | : : : : :
Dd 660 NYNNAVDSANGVTNATSNPNMDANAINOITATQVTSNKNAICPSTNT TQNVOM-----NYV 714

QY	221	ADVTPVNYKARRTVWNGDEPSSRILANTNSITDYSWISLA-----GTNFKY 268
Dh	215	ATDCRMNYKARRTVWNGDEPSSRILANTNSITDYSWISLA-----GTNFKY 268

QY 269 QFSFSNVPSTCYLYFPYKLVKAADA-----NNVGLQYKLN-----NGNVQ 309

QY 310 QVEFATSTAN-----NTTANTPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNHN- 362

831 LAEAKAAARQNLGTLNHIITNAQRTALEGQINQATTV-----DGVNTVKTNANTLDGAMNS 885

Qy	363	-----KVAPMIGNIYLSNEN	378
		: :: :	
Db	886	LQGAINDKDATLRNQNYLDAES	908

RESULT 9

JC6009

surface-located membrane protein lmp3 precursor - Mycoplasma hominis

C:Species: Mycoplasma hominis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: JC6009

R:Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.

J. Bacteriol. 178, 2775-2784, 1996

A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene sys

A:Reference number: JC6009; MUID:56213016

A:Accession: JC6009

A:Molecule type: DNA

A:Residues: 1-1302 <IAD>

A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336

C:Genetics:

A:Gene: lmp3
 A:Genetic code: SGC3
 C:Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homolo
 C:Keywords: duplication; membrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
 F:937-992/Domain: tetratricopeptide repeat homology <TT1>
 F:933-1026/Domain: tetratricopeptide repeat homology <TT2>
 F:1089-1120/Domain: tetratricopeptide repeat homology <TT3>
 F:1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match	8.7%	Score 173;	DB 1;	Length 1302;
Best Local Similarity	22.8%;	Pred. No. 0.079;		
Matches 94;	Conservative	67;	Mismatches 163;	Indels 88;
			Gaps 18;	

Qy	4	ITKKDANPNNGQ-----TQLEAARMELTOLINAKAMT	LASLDQDYAKYEASLSAYSEATV	59
				::
				::
Db	864	ITKKLETFNKDKDVKFKLEQTRKDIDEFN-----TNKTPNDY	STLISELFSKRDNSKI	919

Qy	60	NNLNATLEQLKMAKTNLESAINQANTDKTTFDN-----EHPN-----LVEAYKAL 105 : - :: : : :
Db	920	TNSSKS--DIETANTEUKQALAKANTDKDQADNLARSTKEQLNKSISSANTLLAKLTDR 977 : - :: : : :

Qy	106	KTTLEQRATNLEGLSSTAYNQI-RNNLVLDYINKASSL-----ITKTLDPINGGTL	158
		: :	
Db	978	DNTIQQAKTELEKEVQKQAVASNTASQMSAKSLDAKVTETKLTETFNKDKDYKR	1037

```
QY      159 ETTANKNNITLISTINEOKTNADALSNSFIKKVIONNEOSFVGFTNANVPQSNYSFVA 218  
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      1038 ELEQTRKIDIDEFINT-NKTNPNTSLISELTSK--RDSKSITNSSNKSDIFETANEL-- 1092
```

```

QY 219 PSADVTPVNYKYARRTVVNGDEPSSRIILANTNSITDVSWIYSLAGNTKYQSFSSNYGPS 278
                                     : || | || : || : || : || :
Db 1093 -----KOALAKAN--TDKACADNTA--PSTKEOTNVESTSSAN 1125

```

Qy 279 TGLYFPYKLVRAADANVGLQYKLNNGVQVEFATSTSANNTTANPTPAYDEI -- KVA 336

Db 1126 T-----IIAKITDKDNMTTQQAKTE--TFKVFQAFANQAVASNNVVMQCGGCGGCTDNKMG 1172

QY 337 KIVLSGLRGQNTIELSVPTGEGNNKVPAMIGNI--YLSSNENNADKIPCY 386
: : : : : : : : : : : : : : : :
Db 1177 EITPKLETNNKOK-----EAKFNEUKRGQIQEFINTNNK---PNY 1216

RESULT 10
AG2422

C; Species: *Anabaena* sp.

A; Note: *Anabaena* sp. (s
C: Date: 14-Dec-2001 #s

C;Accession: AG2422
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; I...

K; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

```

QY      8  DANPNNGQTGLE-----AARMELTDLINAKAMTLASLDYAKRIEASLSAYSEAE  57
Db      1186  DQAKTTCETSIDQVTPVKKATARNETAILNNKLEIQIATPDATDEEQAADA---EAN  1243
QY      58  TVNNLN-----ATLEQLKWKATNLESAINQANTDKTTFDNEHPNLYEAYKALKTTL--  109
Db      1244  TENGKANQALSAATTNQAQVDEAKANEAALIN-AVTPKVKKKQAQKDEIQLOQTQTNNVIN  1302
QY      110  -EQRATNLEGLS-----STAYNGIRNNL-----VDLYNKASSLTTKTIDPLNGQTLL  155
Db      1303  NDQONATNEEEKAAIQOLATAVTDAKNNITAAATDDNGVDYAKDAGKNISIQTOP---ATAV  1359

```

```

Db      1360 KSNAKNEVDQAVTTQNOAIDNTTGATTEERNAAKDL-----VLKAKRAYQDIL---N 1409
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
QY      209 VQPSNYSEFAVSADVTPVNYKYARRTWGNGDEPSSRILANTNSTIDVSWIYSLA---GT 264
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      1410 AQTN-----DVTOI-----KQAVADIQGTADTTINDVAKDELAT 1446
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
QY      265 NTKYQFSFSGPSTGYLYFPYKLVKAADAN-----NVGLQYKLNNGNVQVVEFATST 317
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      1447 KANEQKAL-----IAQTADATTEKEQAQOVDAQLTQGN-QNIENAAQSI 1490
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
QY      318 SANNTTA-NTPPAVDIEKVAIKVLGSRFGQNTIELSVPTGEGNNKVKAPMIGNIYLSN 376
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      1491 DVNTAKDNAIQALDPIQASTDVKTNAR-----AELLTEM-----QNKITELNNNETTNE 1541
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
QY      377 ENNAD 381
      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Db      1542 EKGND 1546

RESULT 12
C89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Sema, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaiko, C.; Selander, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus
A:Reference number: A89758; MUID:21311952; PMID:11418146

```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <GR>
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BABA2528.1; GSPDB:GN001
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhB

Query Match      8.2%; Score 163.5; DB 2; Length 3890;
Best Local Similarity 24.4%; Pred. No. 1.1;
Matches 94; Conservative 54; Mismatches 153; Indels 85; Gaps 3

QY 16 TQLEAARMELTDLINAKAMTLASLDYAKTEASLSSAYSEA-ETVNNNLNATLEQLKMAK 74
Db 3528 TDLNTAMGNLQGAINDPQTILNS-QNYQDATPSKKTATYNAVQAQAKDILNKSQNK-TK 3595
QY 75 TNLESAINQANTOKTTTDFNDEHPNLVEAYKALKTTLEQARNLEGLSSTAYNQIRNNLVDL 134
Db 3586 DQVTEAMNQVNSAKNNLDG-----TRLLDQAKQATQKQQLANNMTHLTAAQTNTLTNQ---- 3636

```

```

Db 3637 -----INSGTTVAGVHTVQSNANTLDQAMNTLRQSIANDATKASEDYVDA 3682
Qy 190 ---KKVTIONNEQSVGFTNANVQPS-NYGEVAFSADVTPVNTKYKARTVWNGDE---- 240

```

Db 3683 NNDKOTAYNNAAVAAETIINANSNPNPSPITQKAE--QVN-----SSKTLALNGDENLATA 3737
QY 241 -PSSRIILANT-NSITD-----VSWIYS-----LACTNKKYQSPSNYGPSTGYLYFPFKLV 289
Db 3738 KQNAKTYLNTLTSTDAQKNLISQISSATRVSGVDIVKQNA-----QHL 3782
QY 290 KAADANNVGLQYKLNNGVQVFEATSTSNNTTANTTPAVIDE-IKVAKIVL-----SGLRF 345
Db 3783 DOAMAN-----LQNGLNNSQVKSSEKRYADATNKQEQYDNDATTAAILNKSTGPT 3835
QY 346 GQNTIELS---VPTGEGNMKNKVPAMI 368
Db 3836 AQNAVEAALQRVNTAKDALNGDAKLI 3861
RESULT 13
S41539
fibrinogen-binding protein - Staphylococcus aureus
N:Alternate names: clumping factor
C:Species: Staphylococcus aureus
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S41539; S36630
R:McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*
A:Reference number: S41539; MUID:94224142
A:Accession: S41539
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-933 <MCD>
A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526
Query Match 8.08; Score 159; DB 2; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.29;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;
QY 2 MSITFKDANPNNGTOL--EAARMELDLINAKAMTLASLDQYAKIESLSSAYSEATV 59
Db 73 VSDTKTSNTNNGTSVAQNPAAQQTQSSSTNATT-----EETPVVTEATT 120
QY 60 NNINLATLEOLKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLEQATNLEGL 119
Db 121 TTNOANTPATTSSTNAEELVNG-TSNETTFND--INTVSSVNS-----PQNSTNAEV 172
QY 120 SST-----AYNQRLNVLNLYNK-----ASSLITKLDPLNGGTL 155
Db 173 STTDSTTEATPSNESAPOSTDASNKDVYVQAVNTSAPRMRAFSLAAVAADAPAACTDI 232
QY 156 DSNEITTANKNINNTLTINEQ---KTN-ADALSNSFIK---KVIONNEQSFVGTFTN 206
Db 233 -TNQLTNVTVGIDSGTVYPHQAGYVKNLNGFSPNSAVAGDFTKIIVPKELNNGVST 291
QY 207 ANVQPSNYSFVAFSAVTPVNYKYARTVWNGDEPSSRIILANTNSITDVSWIYSLAG-TN 265
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSDGNVITFTDYVN 324
QY 266 TKQFSPSNYGPSTGYLYFPFKLVKAADANNVGLQYKLNNGVQVFEATSTSNNTAN 325
Db 325 TKDDVKATLMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTAN 362
QY 326 PTPAVDEIKVAKI-----VLSGLRFGQNTIELSVPTGEGNMKNKVPAMI-GNIIYSSNE 377
Db 363 KTVLVDEYKYGKPNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVAPVLTGMLKPNTDS 422
QY 378 N 378
Db 423 N 423

RESULT 14
B71605
hypothetical protein PFB0850c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71605
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.
A:Reference number: A71600; MUID:99021743
A:Accession: B71605
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-807 <GAR>
A:Cross-references: GB:AE001420; GB:AE001362; NID:g3845287; PIDN:AA71959.1; PID:g384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0850c
Query Match 7.88; Score 156; DB 2; Length 807;
Best Local Similarity 23.8%; Pred. No. 0.35;
Matches 73; Conservative 48; Mismatches 124; Indels 62; Gaps 14;
QY 59 VNNLNATLEOLKMAKT-----NLESAINQA--NTDK-----TTFDNEHPNLVEAYKALKT 107
Db 468 INNISTV--OLKMNININFIHESINQOHNNTFFKNNDTNEFETNKKKKKKKEK 525
QY 108 TLEQATNLEGLSTAYNQIRNNILVDLYNKASSLITKLDPLNGGTLDSNEITANKNI 167
Db 526 NIHFNNNNNNNNKCLYKINQ--DHNNSIINTNQNFHDI-----NNVKTEQNL 574
QY 168 ---NNTLSTINEQKTADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVT 224
Db 575 QKHNRKMSQVSKSNKNNKNNSHLKKOININ-----TNNMNMKNNSHISKNVIVD 626
QY 225 PVNYK-----YARTVWNGDEPSS-----RIILANTNSITDVSWIYSLAGTNTKYQ-ESFSN 274
Db 627 DNKLSSHADNSNEIVTKGKKKTKNKKKNNINSVNNVNNNINSNNNIIISNN 686
QY 275 YGPSTGYLYFPFKLVKAADANNVGLQYK-----LNGNVQVQVEFATSTAN 320
Db 687 VNNNNMYPNPVNIQ-KDSDNIALYNNKPNIDFNFNQNLNHNHNNIQQNNITNNVNLN 745
QY 321 N--TTAN 325
Db 746 NNLTTSN 752
RESULT 15
A86827
hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL140
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Bolet, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
A:Reference number: A86825; MUID:Z1235186; PMID:11337471
A:Accession: A86827
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqfG
Query Match 7.88; Score 155; DB 2; Length 1072;
Best Local Similarity 21.0%; Pred. No. 0.59;
Matches 97; Conservative 76; Mismatches 154; Indels 136; Gaps 20;
QY 8 DAPNPNQTOLEAARMELTDLINAKAMTLASLDQYAKIESLSSAYSE-AETVNNNL--- 63

```
Db 430 ESNVSDQTSSEAS-TNSNSSISLSPNISSTSD--SESATNSSDFSNVAEVANNSLASV 486
Qy 64 -NATLEQLKMAKTULEGAINQANTDKTFDNEHNLVEAYKALKYTTLEQRATN----- 115
Db 487 NNSSSVLSSTSTADNLGINQSGSDNLTKDSSEISTSGAFLSSNQTSSEASTNSNSSISL 546
Qy 116 -----LEG-LSSTAYNOI-----RNVLADLYNKASSLI--TKTLDPL-----NGG 152
Db 547 SPSNISSTSVLESTTSSNFSNVAEVANNSLASVNNSSSVLSSTSTADNLEINQFGSDN 606
Qy 153 TLLDSNEITTANKNINNTLSTINQKTNADALSNSFIKKVIQNNQSFVGFFTNANVQPS 212
Db 607 LTKDSSEISTSG-----AFLSSNQTSSEASSNS-----MSSINSPSLSLTNSSESAT 655
Qy 213 NYSFVAFSADVTPVNYKYARTVWNGDEPPSRIL-----ANTNSITDVSWIYISLAG 263
Db 656 NQS---NSSEATKVDN-----NSSTHSSNILLNSGSDSDSDSDSDSSNLSNLSNPN 704
Qy 264 TNTKYQFSFSGYGLYPYKLVKAADANNVGLQYKLNNGNVQVQEFATSTSANNT- 322
Db 705 LETNQTS-----SKPSEYNNI-----SENPKYSSNSNVQENSTD 740
Qy 323 ---TANPTPAYDE-----IKVAKIVLGLRFGONTIE---LSVPTGEG 359
Db 741 HEMSTNPKSSISPISTSSSQKESQSNLLNTTEGINNPITFNSSSENSEASAILTSYS 800
Qy 360 NMKVAPMIGNIYLS-----SNENNADKI 383
Db 801 NNSSESSETGCLYISNEAQRDNGSEISHSLPSSNSNENNVSII 843
```

Search completed: June 12, 2002, 10:49:47
Job time: 205 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:45:47 ; Search time 22.56 Seconds
(without alignments)
425.499 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSITKDPNNGQIQLEA.....SSNNADKIPGVRPCTFL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pap:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pap:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pap:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pap:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pap:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	95.6	615	2	US-08-525-742-10
2	1856	92.9	610	2	US-08-525-742-8
3	1638	82.0	368	2	US-08-525-742-4
4	1612	80.7	368	1	US-08-185-851A-4
5	806	40.3	661	2	US-08-525-742-2
6	159	8.0	933	4	US-08-293-728-2
7	159	8.0	933	4	US-09-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1073	4	US-09-541-782-6
10	132.5	6.6	1002	4	US-09-268-347-24
11	131.5	6.6	1164	4	US-08-923-992A-2
12	131	6.6	1004	4	US-09-268-347-30
13	131	6.6	1104	4	US-08-923-992A-4
14	129	6.5	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	122.5	6.1	1536	1	US-08-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	2	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2

Patent No. 5352450

28	120.5	6.0	746	5	PCT-US95-10509-2	Sequence 2, Appli
29	120.5	6.0	1104	4	US-09-268-347-28	Sequence 28, Appl
30	120.5	6.0	1104	4	US-09-268-347-34	Sequence 34, Appl
31	119.5	6.0	984	1	US-08-242-932-2	Sequence 2, Appli
32	119.5	6.0	984	1	US-08-714-481-2	Sequence 2, Appli
33	119.5	6.0	984	5	PCT-US95-06111-2	Sequence 2, Appli
34	119.5	6.0	2314	4	US-09-268-347-49	Sequence 49, Appl
35	119	6.0	2354	4	US-09-268-347-47	Sequence 47, Appl
36	118.5	5.9	866	2	US-08-483-101-4	Sequence 4, Appli
37	118	5.9	1612	1	US-08-169-927-2	Sequence 2, Appli
38	117.5	5.9	889	4	US-09-336-447A-15	Sequence 15, Appl
39	117.5	5.9	1912	1	US-08-409-995-4	Sequence 4, Appli
40	117.5	5.9	1912	3	US-08-685-467-4	Sequence 4, Appli
41	117	5.9	2353	4	US-09-377-155-33	Sequence 33, Appl
42	117	5.9	2353	4	US-08-913-942-4	Sequence 4, Appli
43	117	5.9	2353	4	US-09-669-974-33	Sequence 33, Appl
44	116	5.8	941	4	US-09-336-447A-9	Sequence 9, Appli
45	116	5.8	1338	2	US-08-728-470-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hiroko

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigemi

; APPLICANT: Takahashi, Kiyohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; ADDRESSEE: NAUGHTON

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-074139

; FILING DATE: 31-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-245625

; FILING DATE: 30-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/00541

; FILING DATE: 31-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mclelland, Le-Nhung

; REGISTRATION NUMBER: 31,541

; REFERENCE/DOCKET NUMBER: 950811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-659-2930

```

; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-10

Query Match          95.6%; Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%; Pred. No. 3.8e-132;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITTKDANPNQGTQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60
Db 27 CMSITTKDANPNQGTQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 86
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 121 STAYNQIRNLDVLYNKASSLIKTLDPLNGGTLDSNEITTKANKNINNTLSTINEQKN 180
Db 147 STAYNQIRNLDVLYNKASSLIKTLDPLNGGTLDSNEITTKANKNINNTLSTINEQKN 206
QY 181 ADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240
Db 207 ADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266
QY 241 PSSRLANTNSITDVSWIYSLAGTNTKYQSFSGNYGPGTGYLYFPYKLVKAADANNVGLQ 300
Db 267 PSSRLANTNSITDVSWIYSLAGTNTKYQSFSGNYGPGTGYLYFPYKLVKAADANNVGLQ 326
QY 301 YKLNGNVQVEFATSTANNTANPTTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
Db 327 YKLNGNVQVEFATSTANNTANPTTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
QY 361 MNKVAPMIGNIYSSNENNAKDI 383
Db 387 MNKVAPMIGNIYSSNENNAKDI 409

RESULT 2
US-08-525-742-8
; Sequence 8, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Sasaki, Sakiko
; APPLICANT: Ohkawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemichi
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; CITY: Washington
; STATE: DC
; STREET: 1725 K Street, Suite 1000
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-8

Query Match          92.9%; Score 1856; DB 2; Length 610;
Best Local Similarity 95.3%; Pred. No. 3.4e-128;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 CMSITTKDANPNQGTQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60
Db 27 CMSITTKDANPNQGTQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 86
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 146
QY 121 STAYNQIRNLDVLYNKASSLIKTLDPLNGGTLDSNEITTKANKNINNTLSTINEQKN 180
Db 147 STAYNQIRNLDVLYNKASSLIKTLDPLNGGTLDSNEITTKANKNINNTLSTINEQKN 206
QY 181 ADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240
Db 207 ADALNSFIKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266
QY 241 PSSRLANTNSITDVSWIYSLAGTNTKYQSFSGNYGPGTGYLYFPYKLVKAADANNVGLQ 300
Db 267 PSSRLANTNSITDVSWIYSLAGTNTKYQSFSGNYGPGTGYLYFPYKLVKAADANNVGLQ 326
QY 301 YKLNGNVQVEFATSTANNTANPTTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
Db 327 YKLNGNVQVEFATSTANNTANPTTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
QY 361 MNKVAPMIGNIYSSNENNAKDI 382
Db 387 MNKVAPMIGNIYSSNENNAKDI 408

RESULT 3
US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Sasaki, Sakiko
; APPLICANT: Ohkawa, Ikuroh
; APPLICANT: Funato, Hirono

```


APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemitsu
APPLICANT: Takahashi, Kiyochiko
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESS: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeLland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-4

Query Match 82.0%; Score 1638; DB 2; Length 368;
Best Local Similarity 95.5%; Pred. No. 1.5e-112;
Matches 322; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 CMSITKDDANPNNGQQLAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
DB 27 CMSITKDDANPNNGQQLAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 86
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 121 STAYNOIRNRLVDLYNKASLIPTKLDPLNGGTLSDSNETTANKNINNTLSTINEQKTN 180
DB 147 STAYNOIRNRLVDLYNKASLIPTKLDPLNGGTLSDSNETTANKNINNTLSTINEQKTN 206
QY 181 ADALNSFKIKVYQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVPYKYARVTWNGDE 240
DB 207 ADALNSFKIKVYQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVPYKYARVTWNGDE 266
QY 241 PSSRILANTNSITDVSNIWISLACTNTKYQSFNSYGFSTGILYFPYKLYKAADANNVGLQ 300
DB 267 PSSRILANTNSITDVSNIWISLACTNTKYQSFNSYGFSTGILYFPYKLYKAADANNVGLQ 326

QY 301 YKLNNGNVQOQVEFATSTSANNTTANPTPAVDEIKVAK 337
DB 327 YKLNNGNVQOQVEFATSTSANNTTANPTQOLMLRLKLLK 363
RESULT 4
US-08-185-851A-4
Sequence 4, Application US/08185851A
Patent No. 5489430
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Fujisawa, Ayumi
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemitsu
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
TITLE OF INVENTION: Well As Vaccines Utilizing the Same
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Armstrong, Westerman, Hattori, McLeLland &
ADDRESS: Naughton
STREET: 1725 K Street, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
SOFTWARE: ASCII from Word Perfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,851A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-887-0357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-185-851A-4

Query Match 80.7%; Score 1612; DB 1; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.2e-110;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY 1 CMSITKDDANPNNGQQLAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
DB 27 CMSITKDDANPNNGQQLAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 86
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 121 STAYNOIRNRLVDLYNKASLIPTKLDPLNGGTLSDSNETTANKNINNTLSTINEQKTN 180
DB 147 STAYNOIRNRLVDLYNKASLIPTKLDPLNGGTLSDSNETTANKNINNTLSTINEQKTN 206
QY 181 ADALNSFKIKVYQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVPYKYARVTWNGDE 240
DB 207 ADALNSFKIKVYQNNNEQSFVGTFTNANVQPSNYSFVAFSADVTVPYKYARVTWNGDE 266

Qy	241	PSSRLANTNSIITDVSWIYSLAGTNTKYQFSFSNYGPTSGYLYFPFKLVKAADANNVGLQ	300
Db	267	PSSRLANTNSIITDVSWIYSLAGTNTKYQFSFSNYGPTSGYLYFPFKLVKAADANNVGLQ	326
Qy	301	YKLNNGVQVQVEPATSTANNTTANPTPAVDEIKVAK	337
Db	327	YKLNNGVQVQVEPATSTANNTTANPTQQLMRKLKLLK	363

RESULT 5
US-08-525-742-2
; Sequence 2, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroko
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemichi
; APPLICANT: Takahashi, Kiyoohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeiland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-2

```

Query Match 40.3%; Score 806; DB 2; Length 661;
Best Local Similarity 42.8%; Pred. No. 2.8e-51;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

[illegible]

```

RESULT 6
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S aureus F
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293-728-2
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

```

Query Match 8.0%; Score 159; DB 3; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.001;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

Qy	2	MSITKDKDANPNNGOTOL--EAARWEITDLINAKAMTILASLDQYAKTEASLSAYSEAEV 59
		: : : : : : : : : : : : : : :
Db	73	VSDTKTSNTNNGETSVQAQOETTQSSSTNATT-----EETPVTGKATTT 120
Qy	60	NNNLNATLEQLKAKTNLESAINQANTDKTTFNEHPNLVEAYKALKTLEQRATNLEGL 119
		: : : : : : : : : : : : : : : : : :
Db	121	TTNQANTPATQSSNTNAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
Qy	120	SST-----AYNQIRNLNLDLYNK-----ASSLIYKTLDPLENGCTLL 155
		: : : : : : : : : : : : : : : : : :
Db	173	STTDQTEATPNNSESPOSTDAKNQVVVQNAVNTSAPRMRAFSLAAVAADAPRAQDIT 232

	Qy	207	ANVOPSNYSFVFASADVTVPWNYKARTVWGDEPSSRIILANTNSTDVSNIYSLAG-TN 265 : :
	Dd	292	KAVPP-----IMAGDQ----VLANGVIDSDGNVIYTFTDYVN 324 : :
	Qy	266	TKYQFSFNPSGPGSYLYFFPKLVKAADANNGLQYKLNNGVNQVEEATSTSANNTTAN 325 : :
	Dd	325	TKDDVKATLTMPA--YI-----DPNV-----KKTGNV-----TLATIGISTTAN 362 : :
	Qy	326	PTPAVDIEIKVAKI-----VLSGRFGQNITIELSVPTGEKNMKVAPMI-GNIYLSSNE 377 : :
	Dd	363	KTVLVDYEKGYGKFNLISIKGTIDDKTNNTYRQTIVVNPSPGDVNAVPLTGNLKPENTDS 422 : :
	Qy	378	N 378
	Dd	423	N 423
		RESULT 7	
		US-09-421-868-2	
		; Sequence 2, Application US/09421868	
		; Patent No. 6177084	
		; GENERAL INFORMATION:	
		; APPLICANT: Foster, Timothy J.	
		; APPLICANT: McDevitt, Damien L.	
		; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene	
		; FILE REFERENCE: 05344.105011	
		; CURRENT APPLICATION NUMBER: US/09/421,868	
		; CURRENT FILING DATE: 1999-10-19	
		; PRIOR APPLICATION NUMBER: 08/293,728	
		; PRIOR FILING DATE: 1994-08-22	
		; NUMBER OF SEQ ID NOS: 20	
		; SOFTWARE: Patentin Ver. 2.0	
		; SEQ ID NO 2	
		; LENGTH: 933	
		; TYPE: PRT	
		; ORGANISM: Staphylococcus aureus	
		US-09-421-868-2	

```

Query Match      8.0%; Score 159; DB 4; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.001;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY          2 MSITKDKANPNNGOTQL--EAARMELTDLINAKAWTLASLDYAKIEASLSAYSEAEV 59
              : || : | ||| : : : : : : : : : : : : : : : : : : : : : :
DB          73 VSDTKTSNTNNGTSSVAQNPAQQTQSSTNATT-----EETPVTEATTT 120

QY          60 NNNLNATLEQLKMAKTNLNESAINQANTDKTTFDNEHPNLVEAYKALKTTLQRATNLBGL 119
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB          121 TTNQANTPATTTQSSNTNAEELVNQ-TSNETTFND--TNTVSVNS-----PQNSTNAENV 172

QY          120 SST-----ANYQIRNNLVLYNK-----ASSLTKTLDPLNGTGILL 155
              : || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB          173 STTQDTSTEATPPSNESAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAGTDI 232

QY          156 DSNETITANKINNITLSTINEQ----KTN-ADALSNSPIK----KVIONNQSFVGTPFN 206
              : || : | : : : : : : : : : : : : : : : : : : : : : : : : : :
DB          233 -TNQLTNVTWVIDSGTIVYPHQAGYVKLYNCFSPNSAVKGDTFKIYVPKEULNGVYST 291

QY          207 ANVOGPSNYSFVASDVTPVNYKYARRTVWNQBDPSSRILANTSITDVSWIYSLAG--TN 265
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB          292 AKVPP-----IMAGDQ-----VLANGVIDSGNVITYFTDYVN 324

QY          266 TKYQFSPSNYGPGTYGILFFPKYLKAADANNVGLQKLNNGNVOOQEATTSANNYTAN 325
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB          325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLTAGIGSTTAN 362

QY          326 PTPAVDEIKVAKI-----VLSLRFGONTIELSVPTGEGNMKNVAPMI-GNTVLSNSE 377
              : | || | || : : : : : : : : : : : : : : : : : : : : : :
DB          363 KTVLYDYEKEYGKFYNLSIKGTIDIDIKTNNTYROTIVYNPSSGDNVIAPLVTLGNLKUPNIDS 422

QY          378 N 378

```

```

Db      423 N 423

RESULT      8
US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

```

Query Match	6.8%;	Score 136;	DB 4;	Length 1098;
Best Local Similarity	22.7%;	pred. No. 0.063;		
Matches	84;	Conservative 51;	Mismatches 159;	Indels 76; Gaps 14;
Qy	4	ITKRDANPNNGOTOLEAARMELTDLINA-KAMTILASLDQYAKIESLSSAYSE-----AE 57		
Db	192	IRKQAQPDKKEDAEVKVREELGKLFSTKAGLQDEIQEHVKKETSSEENTQKYDEHYAN 251		
Qy	58	TVNNLNLAITBOLKMAKTNLESAINQANTDKTTTDFNEHPLN-----VEAYKALK 106		
Db	252	SLQNLAKSLELDKAITN-----EQATQVKNOFLENAQKLKEIOTPLKETNVKLYRAMS 306		
Qy	107	TTLQORATNLEGLSTAYNQIRNNLVLYNKASSLITKTLDPLAGGTLLDSNEITTANKN 166		
Db	307	ESLEQVEKELKHNEA-----NLEDIVAKSKEIVREYEGKLAQSKNLP--ELKQLEEE 357		
Qy	167	INNVLSTINEQ-----KTNDALSNSFIKKVI---QNNQSFVGTFTTNANVPQSNYSFVA 218		
Db	358	AHSLKQVQVEDFRKKFKTSEQVTPKPKRLKRDLAANENNQQKI-----ELTVSPENI--- 408		
Qy	219	FSADVTPVNYKYARRTYVWNGDEPSSRIILANTNSITDYSWIYSLAGTNTKKQFSPS----- 273		
Db	409	-----TVTEGDEVFTVAKSDSKTTLLDFSDLL-----TKYNPVSVDRIST 449		
Qy	274	NYGPSTG---YLYPPYKLVKAADANNVGLQYKLNNGNVQVQFEATSTSANNTTANP--TPA 329		

```
Db 450 NYKTDNHKIAETIKNLKINESQVTLAKDDSGNNVEKTFITVQKKEKQVPTPE 509
QY 330 VDEIKVAKIV 339
Db 510 QKDSKTEEKV 519

RESULT 9
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6

Query Match 6.8%; Score 135.5; DB 4; Length 1073;
Best Local Similarity 20.9%; Pred. No. 0.066;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 13 NGQ-TQLEAARMELTDLINAKA-MTLASL-----QDYAKIEASISAYSEATVNNLNAT 66
Db 611 NGYFTLNDNFASMEELLNTHSNQLLISMTKITEHFQSLDEALQASCSAVPNSSLDLI 670

QY 67 LEQLKMAKTLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEQRATNLEG 118
Db 671 VSELKDSKNSLLDALSHLQDISMSQKLGNGISSELIELQKDKESYROLVQELRSLYN 730

QY 119 LSSTAYNQINNNLDVLYNKASSLITTKTLDPLNGGTLLDSNEIT-----TA 163
Db 731 LQHTHEESQRELGYVRNDIDALVKTCTSLNDADIILSDYISDQKSPESKQODLIANI 790

QY 164 NKINNTLSPINQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 213
Db 791 GKIVSNFLOQNESLYTKADILHSHLNDTNSNIRKANEIMNRSSEFL-----RNA----- 841

QY 214 YSEVAFSADVTPVNYKYARTVWNG-----DEPSSRLANTNSITDVSWYSLA----- 262
Db 842 -----ASQAEIVGANKERIQTVEGSQLDLSKSKATHSRSRYD-----HCLALAESQKQ 893

QY 263 GTWTKYQ-----FSFSGYSPSTGYLYPPYKLVKAADANNVGL--QYKLNNGNVQ 309
Db 894 GVNLEVQTLDRLLQKVKEHSEDNTEKHQOL---LDLLESLVGNNDLIDSIKTPHTELQ 950

QY 310 QVE-----FATSTSANNT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 350
Db 951 KITDHLVKGTLSLANTHLLGDESCLNLETTIEDTSLVKLETTGDTPTPSKRELPAFPS 1010

QY 351 -----ELSVPTGEGNNKVPAMIGNIYLSNENNADKIPGYRRP 389
Db 1011 WTRDSSLIKETTNLNDLSDKKFVREYITSSNQTNPEPV--YDKP 1052
```

```
RESULT 10
US-09-268-347-24
; Sequence 24, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
```

```
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

Query Match 6.6%; Score 132.5; DB 4; Length 1002;
Best Local Similarity 19.1%; Pred. No. 0.1;
Matches 91; Conservative 58; Mismatches 153; Indels 175; Gaps 21;

QY 13 NGOTQLEAARM-----ELTDLINAKAMTSLAQDYAKIEASLSAYSE-----AET 58
Db 362 NGTNPVKISNVADGTEDTDAVSFK--OLKALQD-KQVTLASNAYANGSGSDADGGKATQT 418

QY 59 VNNLNATLEQLKMAKTNLESA-----INQANTDKTTFDN-----EHPN 97
Db 419 LGNDLNFKPKSTDSSELLINIKAAAGDTVTFPPKGSVQVGGDGKATIQDGAKTTGLVEASE 478

QY 98 LVEAYKAL-----KTL-----EQRATN----- 115
Db 479 LVDSLNLKLGKVGKVGKDGATDGTHTDTLVKSGDKVTLKAGDNLKVKQBCNTFTVTVLRD 538

QY 116 -LEGLSSTAYNQIRNNLDVLYNKASSLITK---TLDP---LNGGTLLDSNEITANKNIN 168
Db 539 ELTGVSVEFKDTENGA---NGASTKIYKDGITITPANDANGAATADAKIKVASDGI- 593

QY 169 NTLSTINEQKTNADALSNSFIKKVIQNNESQSFVGTFTNANVQPSNYSFVAFSADVTPVNY 228
Db 594 -----SAGNKAVKNV-----SGLKKEGDANFN-----LTSSADNLTQY 629

QY 229 KYARRTVWNGDEPS-----SRILANTNSITDVSWIYS-----LAGTNTKYQFSFSNYG 276
Db 630 DNAYGLTNLDEKSKQKQPTTVADNTAAATVGDILRGWVISADKTTGESKEYSAQVRNAN 689

QY 277 PSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKVA 336
Db 690 E-----VKFSGSGINVSGLTLDNGTRETIFELAKDEN----- 722

QY 337 KIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVAPMIGNIYLSNEN 378
Db 723 -----AIAFGSGSKALRDNVTAGTGNVYNAEKSGAFGDPNYIEDKAGSYAFGNDN 774
```

```
RESULT 11
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
```

Query Match 6.6%; Score 131.5; DB 4; Length 1164;
Best Local Similarity 22.8%; Pred. No. 0.15;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;
QY 6 KKDANPNNGQTOLEA-ARWELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 58
DB 230 RQAQAADKEDAEVKVREELGKLFSSFKAGLDQIEQHVKKTSSEENTQKVDEHYANS 289
QY 59 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 107
DB 290 LQNLAKSLEELDKATTN-----EQATQVKQFLENAQKLKEIQPLIKETNVKLYKAMSE 344
QY 108 TLEQRATNLEGLSSPAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITPANKNI 167
DB 345 SLEQVEKELKHNSEA-----NLEDVAKSKEIVREYEGKLNQSNLP--ELKQLEEEA 395
QY 168 NNTLSTINEQ-----KTADALSNSFIKKVI---QNNQSFGVTGTNNANVQPSNYSFVAF 219
DB 396 HSKLQVVEDFKKTKTSEQVTPKRVARDLAANNNOOKI-----ELTVSPENI----- 445
QY 220 SADVTVPVNYKARVTWNGDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFSS-----N 274
DB 446 -----TVYEGEDVKFTVAKSDSKTILDSFLL-----TKYNFSVSDRISTN 487
QY 275 YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSANNTTANP-TPAV 330
DB 488 YKNTDNHKAIEITIKNLKNESQTVTLKAKDDSGNVVEKFTTITVQKKEKQVKTPEQ 547
QY 331 DEIKVAKTV 339
DB 548 KDSKTEEKV 556
RESULT 12
US-09-268-347-30
; Sequence 30, Application US/09268347
; Patent No. 6355182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268, 347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1004
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-30

Query Match 6.6%; Score 131; DB 4; Length 1004;
Best Local Similarity 20.2%; Pred. No. 0.13;
Matches 97; Conservative 51; Mismatches 152; Indels 180; Gaps 22;
QY 13 NGQTOLEARM-----ELTDLINAKAMTWLASLDQYAKIEASLSAYSEA-----BT 58
DB 359 NCTNPVKISNVADGTENTDAVSFK--QLKALQD-KQVTLASNAYANGSGDAGGKIQT 415
QY 59 VNNLNATLEQLKMAKTNLESAINQANTDKTTF-----DN----- 93
DB 416 LSNGLN-----PKFKSTDCGLLNIRKAEENDTVTFTPKKGSVQVGDGKATIQDCAKTTGL 470
QY 94 -EHPNLVEAYKAL-----KTTL-----EQRATN-- 115
DB 471 VEASELVDSLNLKLGKVGKGTGCTGTCTGTHDTLVKSGDKVTLKAGDNLKVKQECTNFT 530
QY 116 -----LEGLSSPAYNOIRNNLDVLYNKASSLITK---TLDPLNGTGLDLSNEITPANKN 166
DB 531 YALKDELTDVKSVEPKOTANGA-----NGASTKITKDGTLITPANGAGAAGA----- 577
QY 167 INNTLSTINEQKTADALSNSFIKKVIQNNQSFVGTGTNNANVQPSNYSFVAFSADVTVP 226
DB 578 --NTANTISVTKDGISA-GNKAVKNV-----SGLKKFGDANFDP-----LNTSSADNLTK 624
QY 227 NYKYARRVTWNGDEPS-----SRILANTNSITDVSIIYSLAGT---NTRYQSFSS 273
DB 625 QYDNAYKGLTNLDEKSGKQTPVADNTAAATVGLRGLGWISADKTKGELNKENYNAQVR 684
QY 274 NYGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSANNTTANP-TPAVDEI 333
DB 695 NANE-----VKFKSGNGINVSCKTLDNGTRETITFELAKDEN----- 720
QY 334 KVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVAPMIGNIYLSNEN 378
DB 721 -----AIAFGSGSKALRDNTVAIGTGNVYNAEKSGAFGDPNIEDKAGSYAFGNDN 772
RESULT 13
US-08-923-992A-4
; Sequence 4, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

Query Match 6.6%; Score 131.5; DB 4; Length 1164;
Best Local Similarity 22.8%; Pred. No. 0.15;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;
QY 6 KKDANPNNGQTOLEA-ARWELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 58
DB 230 RQAQAADKEDAEVKVREELGKLFSSFKAGLDQIEQHVKKTSSEENTQKVDEHYANS 289
QY 59 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 107
DB 290 LQNLAKSLEELDKATTN-----EQATQVKQFLENAQKLKEIQPLIKETNVKLYKAMSE 344
QY 108 TLEQRATNLEGLSSPAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITPANKNI 167
DB 345 SLEQVEKELKHNSEA-----NLEDVAKSKEIVREYEGKLNQSNLP--ELKQLEEEA 395
QY 168 NNTLSTINEQ-----KTADALSNSFIKKVI---QNNQSFGVTGTNNANVQPSNYSFVAF 219
DB 396 HSKLQVVEDFKKTKTSEQVTPKRVARDLAANNNOOKI-----ELTVSPENI----- 445
QY 220 SADVTVPVNYKARVTWNGDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFSS-----N 274
DB 446 -----TVYEGEDVKFTVAKSDSKTILDSFLL-----TKYNFSVSDRISTN 487
QY 275 YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNVQOFEFATSTSANNTTANP-TPAV 330
DB 488 YKNTDNHKAIEITIKNLKNESQTVTLKAKDDSGNVVEKFTTITVQKKEKQVKTPEQ 547
QY 331 DEIKVAKTV 339
DB 548 KDSKTEEKV 556
RESULT 12
US-09-268-347-30
; Sequence 30, Application US/09268347
; Patent No. 6355182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268, 347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1004
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-30

Query Match 6.3%; Score 126.5; DB 4; Length 1128;
Best Local Similarity 22.5%; Pred. No. 0.32;
Matches 83; Conservative 51; Mismatches 158; Indels 77; Gaps 15;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:42:12 ; Search time 55.68 Seconds
(without alignments)
783.980 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSTTKDANPNNGQIQLEA.....SSNENADKIPGRRPCTFL 393

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1998	100.0	456	18	AAW36050
2	1994	99.8	1086	18	AAW36051
3	1910	95.6	615	15	AAW63230
4	1856	92.9	610	15	AAW63229
5	1612	80.7	368	14	AAW44493
6	1612	80.7	368	15	AAW63227
7	1604	80.3	368	16	AAW76955
8	1142	57.2	235	10	AAW93646
9	1142	57.2	235	11	AAW05081
10	1142	57.2	261	11	AAW05082
11	1142	57.2	261	16	AAW99911

12	1117	55.9	261	10	AAW93959
13	806	40.3	661	15	AAW63226
14	806	40.3	661	16	AAW79910
15	713	35.7	647	16	AAW11978
16	700.5	35.1	648	15	AAW56973
17	386.5	19.3	183	10	AAW93649
18	386.5	19.3	183	11	AAW08439
19	284.5	14.2	219	11	AAW93648
20	284.5	14.2	219	11	AAW06438
21	176.5	8.8	6281	22	AAU37403
22	176	8.8	1095	22	AAW83030
23	171	8.6	2086	22	AAU34143
24	171	8.6	5795	22	AAU37017
25	162	8.1	1029	22	AAU34389
26	162	8.1	933	21	AAU37490
27	159	8.0	933	21	AAW58435
28	159	8.0	933	22	AAW69508
29	159	8.0	936	18	AAW89801
30	158	7.9	2434	22	AAU34339
31	156	7.8	807	21	AAW18311
32	154.5	7.7	682	17	AAW95273
33	153.5	7.7	2478	22	AAU34320
34	153.5	7.7	2478	22	AAU37374
35	151.5	7.6	1237	18	AAW55640
36	151.5	7.6	1237	20	AAW11787
37	151	7.6	5024	22	AAW82935
38	145.5	7.3	2437	22	AAU34338
39	145	7.3	1215	22	AAU34412
40	145	7.3	1269	22	AAU37520
41	143.5	7.2	837	22	AAU34387
42	143.5	7.2	875	22	AAU37487
43	143	7.2	3158	22	AAU37018
44	142	7.1	135	15	AAW56974
45	142	7.1	135	16	AAW11979

ALIGNMENTS

RESULT

1

AAW36050

ID AAW36050 standard; Protein; 456 AA.

XX AAW36050;

AC AAW36050;

XX 15-JUL-1998 (first entry)

DT 15-JUL-1998 (first entry)

DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

KW antigen; vaccine; poultry.

XX Chimeric - Marek's disease gammaherpesvirus.

OS Chimeric - Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Region 1..64

FT Region /note= "derived from Marek's disease virus gB protein"

FT Region 65..456

FT Region /note= "derived from M. gallisepticum antigenic protein"

XX WO9736924-A1.

XX 09-OCT-1997.

XX 28-MAR-1997; 97WO-JP01084.

XX 29-MAR-1996; 96JP-0103548.

XX (JAPG) NIPPON ZEON KK.

XX Saito S, Tsuzaki Y, Yanagida N;

XX

DR	WPI; 1997-503046/46.
XX	N-PSDB; AAT96595.
PT	Fusion protein comprising herpes virus outer membrane protein and antigenic polypeptide - for prevention of infection by Mycoplasma gallisepticum, especially in poultry
PT	
XX	Disclosure; Page 16-19; Sipp; Japanese.
PS	
XX	This sequence represents the chimeric protein 40 K-S which comprises a fragment of the Marek's disease virus outer membrane protein gB fused to an antigenic protein from Mycoplasma gallisepticum. The chimeric protein can be used in recombinant live vaccines for prevention of infection by Mycoplasma gallisepticum, especially as the outer membrane protein shows antigenicity in poultry.
XX	
SQ	Sequence 456 AA;
	Query Match 100.0%; Score 1998; DB 18; Length 456;
	Best Local Similarity 100.0%; Pred. No. 2.le-126;
	Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CMSITKDDANPNNGQTGLEAARMELTDLINAKAMTLASLDYAKTEASLSAYSEAEVTN 60
Db	64 cmsitkddanpnngqgleaarmeltdlinakamtlaqlgyakleaslsayseaeetvn 123
QY	61 NNLNATLEQLMAKTNTLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGIS 120
Db	124 nlnnatleqlmakntlesainqantdkttfdnehpniveaykaikktieqrattnelegis 183
QY	121 STAYNQIRNNLVLDVYNKASSLITTKDPLNGGTLSDSNEITTANKNNINTTISTINEQKTN 180
Db	184 staynqiirnnlvldynkasslittkdplnggtildsneittanknninttistineqkten 243
QY	181 ADAIENSFIKKVIONNEGSGFTGTNANVPQSNYSFVAFSADVPVNYKYARRTVWNGDE 240
Db	244 adaiensfikkvionnegsfvgttfnanvpqsnysfvafsadvtpvnykyarrtvwngde 303
QY	241 PSSRILANTNSITDVSWIYSLAGNTKVKQFSNVGPSTGYLYPPYKLKKAADANNVGLO 300
Db	304 psrilanntsitdvswiyslagntkvkfsgnygstgyllypyklvkaadannvlglo 363
QY	301 YKLNGNVQQVEFATSANNTTAPAVDEIKRAKIVLSGLRFQGNTIELSVPTGEGN 360
Db	364 yklngnvqqvefatssannttapavdeikvakivlsgrfgntielsvptgegn 423
QY	361 MNKVAPMIGNIYLSSNENNADKIPGYRRPGTFLL 393
Db	424 mnkvapmigniyllssennackipgyrrpgtfll 456
RESULT	2
AAW36051	
ID	AAW36051 standard; Protein; 1086 AA.
XX	
XX	AAW36051;
DT	15-JUL-1998 (first entry)
XX	
DE	Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX	
KW	Chimeric; Marek's disease virus; outer membrane protein; fusion protein; antigen; vaccine; poultry.
XX	
OS	Chimeric - Marek's disease gammaherpesvirus.
OS	Chimeric - Mycoplasma gallisepticum.
XX	
Key	Location/Qualifiers
FT	Region 1..672
FT	/note= "derived from Marek's disease virus gB protein"
FT	693..1086
FT	/note= "derived from M. gallisepticum antigen"
FT	

QY 241 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 267 pssrillantnsitdvsiiyslsgntkyqfsfnsygpstgylyfpyklvkaadannvg1q 326
 QY 301 YKLNNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 360
 DB 327 yklngnvnvqvefatstannnttanptpavdeikvakivlsgrlrfqntielsvptgern 386
 QY 361 MNKVAPMIGNIYLSNENNAKD 382
 DB 387 mnkvapmignmyitsanaek 408
 RESULT 5
 ID AAR44493 standard; Protein; 368 AA.
 AC AAR44493;
 XX 16-JUN-1994 (first entry)
 XX Mycoplasma gallisepticum 40kD antigen.
 DE Vaccine; mycoplasma infection; poultry; fowl.
 KW Mycoplasma gallisepticum.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to NNN codon in AAQ53419"
 FT Misc-difference 283 /note= "corresponds to NNN codon in AAQ53419"
 FT
 FT
 XX WO9324646-A.
 XX 09-DEC-1993.
 XX 28-MAY-1993; 93WO-JP00715.
 XX 29-MAY-1992; 92JP-0138819.
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
 XX WPI; 1993-405837/50.
 DR N-PSDB; AAQ53419.
 XX Mycoplasma gallisepticum antigen and DNA coding for it - useful for vaccination of fowl against mycoplasma infections
 PT
 PT
 XX Claim 2; Page 23-26; 37pp; Japanese.
 The sequence coding for the 40kDa antigen was obtained by PCR amplification of M.gallisepticum genomic DNA. The antigen reacts with Mycoplasma-immune or Mycoplasma-infected serum and can be used as a vaccine to protect fowl from M.gallisepticum infection.
 XX Sequence 368 AA;
 Query Match 80.7%; Score 1612; DB 14; Length 368;
 Best Local Similarity 95.0%; Pred. NO. 1.4e-100;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 QY 1 CMSITKDPANPNNGOTLEAARMETDLINAKMTLASLDYAKTEASLSYSAETVN 60
 DB 27 cmsitkdpnngotlqearmeitdlinakartlasldyakleaslsysaeatvn 86
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEQRATNLEGLS 120
 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEQRATNLEGLS 146

DB 87 nlnnatleqlkmaaktnlesainqantdkttfdnehpnlveaykalkttleqratinlegla 146
 QY 121 STAYNOIRNNLDVLYNKASSLITKTLDPLNGSTLLDSNEITTTANKNNINNTLSTINQOKTN 180
 DB 147 staynoirnnldvlynasslittktldplngmllsneittvnrlnhntcistineqkn 206
 QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARETVWNGDE 240
 DB 207 adalsnsfikkvionneqsfvgtftnanvqpsnysfvafsadvtpvnykyarittvngde 266
 QY 241 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 267 pssrillantnsitdvsiiyslsgntkyqfsfnsygpstgylyfpyklvkaadannvg1q 326
 QY 301 YKLNNGNVOQVEFATSTANNTTANPTPAVDEIKVAK 337
 DB 327 yklngnvnvqvefatstannnttanptqimrliklik 363
 RESULT 6
 ID AAR63227 standard; Protein; 368 AA.
 AC AAR63227;
 XX 23-JUN-1995 (first entry)
 XX Mycoplasma gallisepticum 40kD antigen.
 DE recombinant avipox virus; live vaccine; mycoplasma 40kD antigen; TTM-1.
 KW Mycoplasma gallisepticum.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to a NNN codon"
 FT Misc-difference 283 /note= "corresponds to a NNN codon"
 FT
 XX WO9423019-A.
 XX 13-OCT-1994.
 XX 31-MAR-1994; 94WO-JP00541.
 XX 31-MAR-1993; 93JP-0074139.
 XX 30-SEP-1993; 93JP-0245625.
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 XX Saeki S, Saitos, Takahashi K;
 XX WPI; 1994-333181/41.
 DR N-PSDB; AAQ77854.
 XX Recombinant avipox virus combining DNA encoding a polypeptide - exhibiting antigenicity of mycoplasma, useful for the production of a live vaccine
 PT
 PT
 XX Claim 4; Page 71-74; 123pp; Japanese.
 The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding for the 40kD antigen of Mycoplasma gallisepticum under the control of a synthetic promoter. A 1300 bp restriction fragment containing the promoter-ORF sequence was excised and was used in the construction of plasmid pNZ7929-R2. This in turn was involved in the construction of a recombinant avipox vector comprising the TTM-1 gene, DNA encoding the signal membrane anchor peptide from Newcastle Disease Virus haemagglutinin neuraminidase and FPV sequences. The recombinant avipox virus is useful as a live vaccine

CC to protect against infection by Mycoplasma gallisepticum.

XX Sequence 368 AA;

Query Match 80.7%; Score 1612; DB 15; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.4e-100;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 CMSITKDDANPNNGQTGLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
DB 27 cmsitkddanpnngqtlqgarmeltdlinakartlasldyakeaslsayseaeetvn 86
QY 61 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
DB 87 nninatleqlkmaiktlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146
QY 121 STAYNOIRNNLDLYNKASSLIKTLDPLNGGTLTLDSEITNTANKNNINNTLSTINEQKTN 180
DB 147 staynqirnnldvlynnassliktldplnggmldseitntvnrnnntlstineqktn 206
QY 181 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
DB 207 adalsnfsikkvionneqsfvgtftnanvqpsnysfvafsadvtpvnykyartvxnge 266
QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNYPGSGTGLYFPYKLVKAADANNVGLQ 300
DB 267 psrirlantnsitdvswiyxylagntkyqfsfnypgstgylfpyklvkaadannvqlq 326
QY 301 YKLNNGNVQVEFATSTSANNTTANPTPAVDIEKVK 337
DB 327 yklnngnvqvfeafatsannttanptqqlmrklk 363

RESULT 7

AAR76955
ID AAR76955 standard; Protein; 368 AA.

XX AAR76955;

DT 12-MAR-1996 (first entry)

XX Mycoplasma gallisepticum antigenic protein TTM-1.

XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Misc-difference 262 /note= "any amino acid"

FT Misc-difference 283 /note= "any amino acid"

XX JP07133295-A.

XX 23-MAY-1995.

XX 27-AUG-1993; 93JP-0213102.

XX 27-AUG-1993; 93JP-0213102.

XX (SHIO) SHIONOGI & CO LTD.

XX WPI; 1995-220782/29.

XX N-PSDB; AAQ94711.

XX A new antigenic protein which reacts with Mycoplasma gallisepticum -
is useful in a component vaccine for use against poultry infected
with M. gallisepticum.

XX Claim 6; Figs 5-6; 33pp; Japanese.

XX

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
CC Infectious diseases in poultry, and as a diagnostic agent for
CC M. gallisepticum.

XX Sequence 368 AA;

Query Match 80.3%; Score 1604; DB 16; Length 368;
Best Local Similarity 94.4%; Pred. No. 4.7e-100;
Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 CMSITKDDANPNNGQTGLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
DB 27 cmsitkddanpnngqtlqgarmeltdlinakartlasldyakeaslsayseaeetvn 86
QY 61 NNINATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
DB 87 nninatleqlkmaiktlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146
QY 121 STAYNOIRNNLDLYNKASSLIKTLDPLNGGTLTLDSEITNTANKNNINNTLSTINEQKTN 180
DB 147 staynqirnnldvlynnassliktldplnggmldseitntvnrnnntlstineqktn 206
QY 181 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
DB 207 adalsnfsikkvionneqsfvgtftnanvqpsnysfvafsadvtpvnykyartvxnge 266
QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNYPGSGTGLYFPYKLVKAADANNVGLQ 300
DB 267 psrirlantnsitdvswiyxylagntkyqfsfnypgstgylfpyklvkaadannvqlq 326
QY 301 YKLNNGNVQVEFATSTSANNTTANPTPAVDIEKVK 337
DB 327 yklnngnvqvfeafatsannttanptqqlmrklk 363

RESULT 8

AAP93646

ID AAP93646 standard; protein; 235 AA.

XX AAP93646;

XX 11-MAY-1990 (first entry)

XX Amino acid sequence of Mycoplasma gallisepticum (MG1) polypeptide.

XX Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0005441.

XX 02-JUN-1988; 88JP-0136343.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.

XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX WPI; 1989-358393/49.

XX N-PSDB; AAN92568.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
vaccines.

XX Disclosure; Fig.1a; 31pp; English.

XX This amino acid sequence of MG1 is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
 XX
 SQ Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 10; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEAEVTN 60
 |||||
 Db 1 cmsitkddanpnngqtqlqaarmeltdlinakartlasldqyakieaslsayseaevtn 60
 |||||
 Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 |||||
 Db 61 nlnnatleqlkmaктnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 120
 |||||
 Qy 121 STAYNQIRNNLVLYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
 |||||
 Db 121 staynqirnnlvlynnasslittkldplngmllidsneittvnrnintlstineqktn 180
 |||||
 Qy 181 ADALNSFIKKVIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPVNYKYARRTV 235
 |||||
 Db 181 adalnsfikkvionneqsfvgtftnanvpqsnysfvafsadvtpvnykyarrtv 235
 |||||

RESULT 9

AAR05081
 ID AAR05081 standard; protein; 235 AA.

XX AC AAR05081;
 XX
 DT 08-OCT-1990 (first entry)
 XX MG-1 antigen.
 XX Mycoplasma gallisepticum; poultry; vaccine.
 XX JP02111795-A.
 XX 24-APR-1990.

XX 02-JUN-1989; 89JP-0136343.
 XX 02-JUN-1989; 89JP-0136343.

XX (JAPG) NIPPON ZEON KK.
 XX (SHIO) SHIONOGI KK.

XX WPI; 1990-169109/22.
 XX N-PSDB; AAQ04686.

XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.

XX Claim 2; Fig 1a; 20pp; Japanese.

XX DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081-2 and AAR06437-41.

XX Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 11; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEAEVTN 60
 |||||
 Db 1 cmsitkddanpnngqtqlqaarmeltdlinakartlasldqyakieaslsayseaevtn 60
 |||||
 Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 |||||
 Db 61 nlnnatleqlkmaктnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 120
 |||||
 Qy 121 STAYNQIRNNLVLYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
 |||||
 Db 121 staynqirnnlvlynnasslittkldplngmllidsneittvnrnintlstineqktn 180
 |||||
 Qy 181 ADALNSFIKKVIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPVNYKYARRTV 235
 |||||
 Db 181 adalnsfikkvionneqsfvgtftnanvpqsnysfvafsadvtpvnykyarrtv 235
 |||||

RESULT 10

AAR05082
 ID AAR05082 standard; protein; 261 AA.

XX AC AAR05082;
 XX
 DT 08-OCT-1990 (first entry)
 XX TMG-1 antigen.
 XX Mycoplasma gallisepticum; poultry; vaccine.
 XX JP02111795-A.
 XX 24-APR-1990.

XX 02-JUN-1989; 89JP-0136343.
 XX 02-JUN-1989; 89JP-0136343.

XX (JAPG) NIPPON ZEON KK.
 XX (SHIO) SHIONOGI KK.

XX WPI; 1990-169109/22.
 XX N-PSDB; AAQ04687.

XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.

XX Claim 2; Fig 2; 20pp; Japanese.

XX DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.

XX Sequence 261 AA;

Query Match 57.2%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 3.4e-69;
 Matches 229; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEAEVTN 60
 |||||
 Db 27 cmsitkddanpnngqtqlqaarmeltdlinakartlasldqyakieaslsayseaevtn 86
 |||||
 Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 |||||
 Db 87 nlnnatleqlkmaктnlesainqantdkttfdnehpnlveaykalkttleqratnleгла 146
 |||||
 Qy 121 STAYNQIRNNLVLYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
 |||||

```

Db 147 staynqirnnldvlynnaslltktdplngmllidsneittvnrnintlstineqktn 206
QY 181 ADALNSFSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 235
Db 207 adalnsfikkvignegsfvgtftnvnqpsnysfvafsadvtvpnykyartv 261

RESULT 11
AAR79911
ID AAR79911 standard; Protein; 261 AA.
AC AAR79911;
XX
XX
DT 19-JUL-1996 (first entry)
DE M.gallisepticum 261 amino acid protein.
KW Detection; probe; primer; PCR; amplification; secretion; lung;
KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX
OS Mycoplasma gallisepticum.
XX
XX JP07236498-A.
XX
XX 12-SEP-1995.
XX
XX 25-FEB-1994; 94JP-0052764.
XX
XX 25-FEB-1994; 94JP-0052764.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX WPI; 1995-347462/45.
XX N-PSDB; AAT04076.
XX
XX Detection of Mycoplasma gallisepticum - for the quick detection,
XX i.e. within one day, of avian chronic respiratory diseases
XX
XX Claim 3; Page 10-11; 11pp; Japanese.
XX
XX This is the amino acid sequence of a 261 amino acid protein encoded
XX by a fragment of the Mycoplasma gallisepticum genome. The encoding
XX sequence and the sequence of AAT04075 (encoding a 661 amino acid
XX protein) can be used to detect M.gallisepticum using probes based on
XX nucleotides 1125-1648 and primers based on nucleotides 449-456, the
XX complement of bases 893-919, 1908-1934 and the complement of bases
XX 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
XX encoding this protein. The method using these sequences is faster i.e. is
XX able to detect M.gallisepticum, which causes avian chronic respiratory
XX diseases, within one day, from avian secretions, washings from the lung,
XX respiratory tract, nasal cavity, etc.
XX
XX Sequence 261 AA;

Query Match 57.2%; Score 1142; DB 16; Length 261;
Best Local Similarity 97.0%; Pred. No. 3.4e-69;
Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 60
Db 27 cmsitkkdanpnngqtlqqaarmeltdlinakartlasldqyaki easlsayseatvn 86
QY 61 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 120
Db 87 nnlnatleqlkma kt nlesalinqandkttf dnehpnlveaykalkttleqratnle gla 146
QY 121 STAYNQIRNNLDVLYNNKASLLTKTLDPLNGGTLSDSNEITTANKNNINTLSTINEQKT 179
Db 147 staynqirnnldvlynnaslltktdplngmllidsneittvnrnintlstineqktn 206
QY 180 ADALNSFSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 235
Db 206 adalnsfikkvignegsfvgtftnvnqpsnysfvafsadvtvpnykyartv 261

RESULT 13
AAR63226
ID AAR63226 standard; Protein; 661 AA.
XX

```

```

QY 181 ADALNSFSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 235
Db 207 adalnsfikkvignegsfvgtftnvnqpsnysfvafsadvtvpnykyartv 261

RESULT 12
AAP93959
ID AAP93959 standard; protein; 261 AA.
XX
XX AAP93959;
XX
XX 11-MAY-1990 (first entry)
DE Amino acid (AA) sequence of TMG-1 polypeptide.
XX
KW Mycoplasma gallisepticum; Poultry vaccine; ss;
XX
OS Mycoplasma gallisepticum.
XX
XX EP345021-A.
XX
XX 06-DEC-1989.
XX
XX 31-MAY-1989; 89EP-0005441.
XX
XX 02-JUN-1988; 88JP-0136343.
XX
XX (JAPG ) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.
XX
XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
XX WPI; 1989-358393/49.
XX N-PSDB; AAN92574.
XX
XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
XX vaccines.
XX
XX Disclosure; Fig.2; 31pp; English.
XX
XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
XX same sequence as that of a polypeptide expressed in Mycoplasma
XX gallisepticum in nature. When the corresponding DNA sequence is inserted
XX into a recombinant vector used to transform a host the antigen protein
XX produced can be used as a vaccine to prevent and diagnose MG infection.
XX
XX Sequence 261 AA;

Query Match 55.9%; Score 1117; DB 10; Length 261;
Best Local Similarity 96.2%; Pred. No. 1.6e-67;
Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 60
Db 27 cmsitkkdanpnngqtlqqaarmeltdlinakartlasldqyaki easlsayseatvn 86
QY 61 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 120
Db 87 nnlnatleqlkma kt nlesalinqandkttf dnehpnlveaykalkttleqratnle gla 146
QY 121 STAYNQIRNNLDVLYNNKASLLTKTLDPLNGGTLSDSNEITTANKNNINTLSTINEQKT 179
Db 147 staynqirnnldvlynnaslltktdplngmllidsneittvnrnintlstineqktn 206
QY 180 ADALNSFSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 235
Db 206 adalnsfikkvignegsfvgtftnvnqpsnysfvafsadvtvpnykyartv 261

RESULT 13
AAR63226
ID AAR63226 standard; Protein; 661 AA.
XX

```

```

AC AAR63226;
XX
DT 23-JUN-1995 (first entry)
XX
DE Mycoplasma gallisepticum antigen (UM-81).
XX
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Protein 1..661
FT /note= "Trp residues correspond to TGA codons"
XX
PN WO9423019-A.
XX
PD 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
PR 31-MAR-1993; 93JP-0074139.
PR 30-SEP-1993; 93JP-0245625.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
PI Saeki S, Saitos, Takahashi K;
XX
DR WPI; 1994-333181/41.
DR N-PSDB; AAQ77853.
XX
PT Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PT of a live vaccine
XX
PS Claim 4; Page 61-65; 123pp; Japanese.
XX
CC A restriction fragment of the insert of M.gallisepticum genomic clone
CC pUM-81 containing an open reading frame was sequenced (AAQ77853). The
CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox
CC virus comprising the coding sequence can be used as a live vaccine to
CC protect against infection by Mycoplasma gallisepticum.
XX
SQ Sequence 661 AA;

Query Match 40.3%; Score 806; DB 15; Length 661;
Best Local Similarity 42.8%; Pred. No. 4.4e-46;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

QY 8 DAMPNNGQTOLEAARMELTD-----LINAKAWTLASLDYAKIEASLSAYSEATVN 60
Db : ||: || : || || || || || || || || || || || || || || || || :
45 ntspdsqggmmnaaakeladakaaltllngetaniasyedyakikseltsayetakavs 104
QY 61 NNLNATLEQLKMAKTINLESAINQANTDKTTFDNEHPLVYAYKALKTTLEQRATNLEGLS 120
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
105 akkgatlinevneaktldaaikkaasakndfdaghslvaynnlketlkeetnldsla 164
QY 121 STAYNQIRNNLDVLYKASSLIYKTLDPPLNGGVLLOSNETTTANKNNINTLSTINEQKTN 180
Db : || || || || || || || || || || || || || || || || || || ||
165 nenyaaairtnalnalyekantivtatidpat-gnipevmstvqanqdnatnslriawkqn 223
QY 181 ADALNSFIKKVQIQNEQSPVGFNTANVQPSNYSFVAESADYTPVNYKYARTVWNGDE 240
Db || : || || || : || : || : || : || : || : || : || : || : ||
224 adhlansflqslvknlttrvdvanqge-qpanysfvgsnvndtpnwnfagrkvvasen 282
QY 241 -----PSSRLANTNSITDVSWIYSLAGTNTKRYQSFNSYNG-PSTGYLYFPYKLVKAAAD 293
Db : || || || || || || || || || || || || || || || || || || ||
283 tplatpaedatqgaalsldvswiyslngaeakytlsfrygaeakytaylyfpyklvkted 342
QY 294 ANNVGLOYLKNGNVOQVEF-----ATSTSAN--NTTANP-----TPAVDEIKVAKIV 339
||| ||| ||| : || : || : || : || : || : || : || : || : || :

```

```

Db 343 --nvglqyklnggdkqinfvqtspasgssdvaaneetmaspaemqasptvddikiakva 400
QY 340 LSGLRFQONTIELSVPTGEGNNKNKVPAMIGNIYLSS-----NENN 379
||: || |||| |||| || |||| |||| || |||| |||| || ||||
Db 401 lsnlkfsntiefsvptg-----kaapmignmyitssnsevnknkiyddlfgnsfnenn 455
QY 380 -----ADKIPGY 386
Db 456 ptavtvdllkgy 467

RESULT 14
AAR79910
ID AAR79910 standard; Protein: 661 AA.
XX
AC AAR79910;
XX
DT 19-JUL-1996 (first entry)
XX
DE M.gallisepticum 661 amino acid protein.
XX
KW Detection; probe; primer; PCR; amplification; secretion; lung;
KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Misc-difference 220 /note= "encoded by TGA"
FT Misc-difference 270 /note= "encoded by TGA"
FT Misc-difference 305 /note= "encoded by TGA"
FT Misc-difference 581 /note= "encoded by TGA"
XX
PN JP07236498-A.
XX
PD 12-SEP-1995.
XX
PF 25-FEB-1994; 94JP-0052764.
XX
PR 25-FEB-1994; 94JP-0052764.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
XX WPI; 1995-347462/45.
XX N-PSDB; AAT04075.
XX
PT Detection of Mycoplasma gallisepticum - for the quick detection,
PT i.e. within one day, of avian chronic respiratory diseases
XX
PS Claim 2; Page 7-10; 11pp; Japanese.
XX
CC This is the amino acid sequence of a 661 amino acid protein encoded
CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
CC sequence and the sequence of AAT04076 (encoding a 261 amino acid
CC protein) can be used to detect M.gallisepticum using probes based on
CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
CC complement of bases 893-919, 1908-1934 and the complement of bases
CC 2184-2210 of the sequence, and a probe based on nucleotides 718-41 of
CC AAT04076. The method using these sequences is faster i.e. is able to
CC detect M.gallisepticum, which causes avian chronic respiratory diseases,
CC within one day, from avian secretions, washings from the lung,
CC respiratory tract, nasal cavity, etc.
XX
SQ Sequence 661 AA;

Query Match 40.3%; Score 806; DB 16; Length 661;
Best Local Similarity 42.8%; Pred. No. 4.4e-46;
Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

```



```

QY      8 DANPNNGOTOLEARWELTD-----LINAKAMTILASLODYAKITPASLSAYSAETVN   60
Db      45 ntupsdgggmnaaakeIadakaalttIngetanlasyedyakIseltsayetakavs   104
QY     61 NNUNLATEQLKMKARTLESAINQANTDKTTFDNEHENLVAYKALKTTTLEQRATNLLEGIS  120
Db    105 aktgatinevneaktIdaaIkkaasakndfdaqhslyeaynnIketiKeektnldsls   164
QY    121 STAYNQIRNNVLVOLYNKASSLIKTLDPLNGCTLLDSNITTANKWNINTLTSTINEQRTN   180
Db    165 nenyaarlnslsyekanvicatIdpat-gnipevmstvqangqditnatrIlawkqn   223
QY    181 ADALSNSFIKKVTONNEQSEVFGTTFNANVOPSNYSFVAFSADVPVNYKYARTVVMGDE   240
Db    224 adnlansfikgsivknltrdvvanqe-qpanysvfgsvnvdtpnwnfaqrkvwasen   282
QY    241 -----PSSRILANTNSITDVSNIYSLAGNTNKYQPSFSNYG-PSNGYLYLFYPKLVKAAD   293
Db    283 tplatpaedatqaasltadvswylngaeakytlisrfygaektaylffpyklvktsd   342
QY    294 ANNVGLQYLKLNNGVNOQVEF-----ATSTAN--NTTANP-----TPAVDEIKVAKIV   339
Db    343 --nvglgdykinggdtkqinfvqtpagsgsdvaaneetmaspaemqsaPTvdvdiakva   400
QY    340 LSGLRFGONTIELSVPTPEGNMKNKAPMTIGNIYLS-----NNNN          379
Db    401 lsnlkfnstiefsvptg----kaapmignmyltsnsnevknkyddlfngsfnnenn   455
QY    380 -----ADKIPGY 386
Db    456 ptavtvdlIkgy 467

```

RESULT	15	
AAW11978		
ID	AAW11978 standard; Protein; 647 AA.	
XX		
XX	AAW11978;	
XX		
DT	23-APR-1997 (first entry)	
XX		
DE	Mycobacterium gallisepticum PMGAL.2 adhesin.	
XX		
KW	Adhesin; pmga; mycoplasma; diagnosis; vaccine; vector;	
KW	respiratory disease; poultry; haemagglutinin.	
XX		
OS	Mycoplasma gallisepticum strain S6.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..25
FT		/label= Sig_peptide
FT		/note= "the signal peptide shows homology to
FT		the PMGAL.3 signal peptide"
XX		
PN	CA2135330-A.	
XX		
PD	11-MAY-1995.	
XX		
PF	08-NOV-1994; 94CA-2135330.	
XX		
PR	20-APR-1994; 94US-0230312.	
PR	10-NOV-1993; 93AU-0050593.	
XX		
PA	(BROW/) BROWNING G F.	
XX		
PI	Browning GF, Glew MD, Markham PE, Walker ID, Whithear KG;	
XX		
DR	WPI: 1995-241027/32.	
DR	N-PSDB; AAT51531.	
XX		
PT	New promoter region from a Mycoplasma gallisepticum adhesin gene -	
PT	useful when coupled to foreign antigen gene, for prodn. of	

PT	multivalent live vaccines, also new probes for detecting Mycoplasma
FT	and manipulating its genome
XX	
PS	Disclosure; Fig 3; 8lpp; English.
XX	
CC	Adhesin pMGAL.2 (AAW11978) and adhesin pMGAL.3 fragment (AAW11979)
CC	are products of gene sequences (see also AA51531) isolated from
CC	Mycoplasma gallisepticum. DNA constructs incorporating the
CC	promoter and/or signal sequences of the pMGa genes can be used
CC	in the prodn. of multivalent live vaccines. The signal peptide
CC	sequence is utilised where attachment of an exogenous antigen
CC	gene to the mycoplasma cell membrane is required.
XX	
SQ	Sequence 647 AA;
Query Match 35.7%; Score 713; DB 16; Length 647;	
Best Local Similarity 41.1%; Pred.No. 7.7e-40;	
Matches 169; Conservative 60; Mismatches 124; Indels 59; Gaps	
Qy	8 DANPNQG-----TGLEARVELTLINAKAWTLASLDYAKIEASLSAYSEATYN 60
Db	: : : : : : : : : :
48	dtnpgdgqgmnaasqelaarnglttcfidskagnlglydyakttgntlkaydaaktvl 107
Qy	61 NNLNATLEQLKMAKTIESAINQANTDKTTFDNEHPNLVEAYKALKTTTLEQRATNLEGIS 120
Db	: : : : : : : : : :
108	dnsssttcqlneaktrietairtaatakqtfddeghaelkvkykelkttiisetatlapya 167
Qy	121 STANQNENNVLYDYNKASSLIKTKLPLNGGTLILDSNETTANKMINNTL--STNEOK 178
Db	: : : : : : : : : :
168	.dagvagikmlhslgydagkaiktctlepegdp-ltasavmmantkiveaikdevlnpqk 226
Qy	179 TN----ADALSNSPIKKVIONNEQSFGVTFTNAVOPSNYSFVAFSADVTPV----- 226
Db	: : : : : : : :
227	enatkldslslsvkitgtvee-----ahnkapanysfygykrwytellldkqvfp 279
Qy	227 NYKYARTPVV-NGDEPSSRIANT-----NSITDVSWIYSLAGNTKYQSFNSYGP 277
Db	: : : : : : : : : : :
280	nwdvaertiftnsdcp--rsaintpadqqmacplsnvswisylagtkytleftyyypg 337
Qy	278 STGYLYFPYKLVLKAADANNVGLOYKLNGNVQQVEATSTSA-----NNTANPTPAVD 332
Db	: : : : : : : : : :
338	stgylfpyklvntsdgvkigleykln-----atpsaitfgsdqmngktcptvd 389
Qy	333 IKVAKIVSLGRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNNENNADKI 383
Db	: : : : : : : : : :
390	invakvtlanlfngskniefsupa-----ekvsrpmigmnylssspwnwki 435
Db	: : : : : : : : : :

Search completed: June 12, 2002, 10:48:38
Job time: 386 sec

PT New promoter region from a Mycoplasma gallisepticum adhesin gene -
PT useful when coupled to foreign antigen gene, for prodn. of

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:58:45 ; Search time 5.42292 Seconds
(without alignments)
2997.887 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFFLIVLYGTN.....STFYLCPPPGVGVIRLEFG 63

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	314	96.3	805	12 Q98Y44	Q98Y44 turkey herp
2	314	96.3	805	12 Q98Y45	Q98Y45 turkey herp
3	314	96.3	865	12 Q83291	Q83291 marek disea
4	314	96.3	865	12 Q8JLW2	Q8JLW2 turkey herp
5	314	96.3	865	12 Q8JLW3	Q8JLW3 turkey herp
6	314	96.3	865	12 Q8JLW4	Q8JLW4 turkey herp
7	184	56.4	864	12 Q69408	Q69408 meleagrid h
8	184	56.4	870	12 Q9DPQ9	Q9dpq9 meleagrid h
9	184	56.4	870	12 Q9ELG4	Q9elg4 meleagrid h
10	171.5	52.6	865	12 Q69406	Q69406 turkey herp
11	171.5	52.6	865	12 Q9PWZ1	Q9pwzi turkey herp
12	107.5	33.0	879	12 Q9ISD0	Q9isd0 canine herp
13	107	32.8	891	12 Q9J052	Q9j052 simian herp
14	107	32.8	891	12 Q66018	Q66018 simian herp
15	106	32.5	948	12 Q90050	Q90050 simian herp
16	105	32.2	881	12 Q84735	Q84735 phocine her

17	100	30.7	868	12 Q98VM4	Q98vm4 human herpe
18	99.5	30.5	943	12 Q86665	Q86665 feline herp
19	96	29.4	885	12 Q69387	Q69387 cercopithe
20	96	29.4	893	12 Q65540	Q65540 baboon herp
21	96	29.4	908	12 Q65538	Q65538 baboon herp
22	95	29.1	903	12 Q69076	Q69076 human herpe
23	95	29.1	904	12 Q9DXE4	Q9dxe4 herpes simp
24	95	29.1	904	12 Q91WU4	Q91wu4 human herpe
25	95	29.1	904	12 Q69526	Q69526 human herpe
26	95	29.1	904	12 Q9QLM8	Q9qlm8 human herpe
27	91	27.9	950	12 Q66678	Q66678 equine herp
28	91	27.9	980	12 Q66682	Q66682 equine herp
29	90	27.6	931	12 Q9QAP6	Q9qap6 rangiferine
30	87.5	26.8	919	12 Q9QAP5	Q9qap5 caprine her
31	87	26.7	894	12 Q9YW92	Q9yw92 macropodid
32	86	26.4	916	12 Q66016	Q66016 cercopithe
33	86	26.4	975	12 Q39275	Q39275 equine herp
34	85.5	26.2	947	12 Q9QAP7	Q9qap7 bovine herp
35	85	26.1	901	12 Q69464	Q69464 herpes simp
36	85	26.1	901	12 Q69465	Q69465 herpes simp
37	85	26.1	902	12 Q69095	Q69095 herpes simp
38	85	26.1	904	12 Q37453	Q37453 herpes simp
39	85	26.1	904	12 Q89920	Q89920 herpes simp
40	85	26.1	904	12 Q91C63	Q91c63 herpes simp
41	84	25.8	906	12 Q91WU3	Q91wu3 pseudorabie
42	84	25.8	920	12 Q8JY98	Q8jy98 herpesvirus
43	83	25.5	912	12 Q8QV01	Q8qv01 buffalo her
44	80	24.5	887	12 Q9YW93	Q9yw93 macropodid
45	78	23.9	929	12 Q9QAP4	Q9qap4 cervid herp

ALIGNMENTS

RESULT 1

Q98Y44 PRELIMINARY; PRT; 805 AA.
 ID Q98Y44
 AC Q98Y44;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Glycoprotein B (Fragment).
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Davidson I.;
 RT "Marek's disease virus - turkey isolate, g8.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF350326; AAK37557.2; .
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; GlycoproteinB; 1.
 DR ProDom; PD000693; GlycoprotB; 1.
 FT NON_TER 805
 SQ SEQUENCE 805 AA; 91115 MW; FC8EB090885E15AA CRC64;

Query Match 96.3%; Score 314; DB 12; Length 805;

Best Local Similarity 100.0%; Pred.No. 6.5e-32;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSSPSTQNTVSREVWSSVQLSEESTFYLCPPPGVGVIRL 60

|||||

Db 1 MHYFRNCIFFLIVLYGTNSSPSTQNTVSREVWSSVQLSEESTFYLCPPPGVGVIRL 60

|||||

QY 61 E 61

|

Db 61 E 61

RESULT 2

Q98Y45

```
ID Q98Y45 PRELIMINARY; PRT; 805 AA.
AC Q98Y45;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B (Fragment).
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson I.;
RT "Marek's disease virus - chicken isolate, gB gene.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF30325; AAK37556.2; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
FT NON_TER 805
SQ SEQUENCE 805 AA; 91111 MW; F88B4410A85E15AA CRC64;

Query Match 96.3%; Score 314; DB 12; Length 805;
Best Local Similarity 100.0%; Pred. No. 6.5e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
QY 61 E 61
DB 61 E 61

RESULT 3
ID Q83291 PRELIMINARY; PRT; 865 AA.
AC Q83291;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein 100 precursor.
GN GB.
OS Marek disease virus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=38013;
RN [1]
RP SEQUENCE OF 1-498 FROM N.A.
RA STRAIN-JW;
RA Sousloparov M.A., Bakhtina M.M., Krendelshtchikov A.V., Babkin I.V.;
RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
complex gp100, gp60, gp48 (p-antigen).";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X91985; CAA63039.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;

Query Match 96.3%; Score 314; DB 12; Length 865;
Best Local Similarity 100.0%; Pred. No. 7e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
QY 61 E 61
DB 61 E 61
```

```
Db 61 E 61

RESULT 4
ID Q8JLW2 PRELIMINARY; PRT; 865 AA.
AC Q8JLW2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-attenuated GA;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (VMDV, VVMDV, VV+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129969; AAM97702.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98020 MW; F4F526114EA02DE5 CRC64;

Query Match 96.3%; Score 314; DB 12; Length 865;
Best Local Similarity 100.0%; Pred. No. 7e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
QY 61 E 61
DB 61 E 61

RESULT 5
ID Q8JLW3 PRELIMINARY; PRT; 865 AA.
AC Q8JLW3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-N;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (VMDV, VVMDV, VV+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129967; AAM97700.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98030 MW; 9008B58B484E114D CRC64;

Query Match 96.3%; Score 314; DB 12; Length 865;
Best Local Similarity 100.0%; Pred. No. 7e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
Db 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 E 61
Db 61 E 61

RESULT 6

Q8JLW4 PRELIMINARY; PRT; 865 AA.
AC Q8JLW4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-X;
RA Parcellis M.S., Shamblin C.E., Dienglewiez R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT Pathotypes of Marek's Disease Viruses (vMDV, vVMDV, vv+MDV): Mutations
RT in the Glycoprotein L-encoding Gene in Some vv+MDVs.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129565; AA097698.1; -;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR Prodom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 98107 MW; 4D2628B5E4DEB28B CRC64;

Query Match 96.3%; Score 314; DB 12; Length 865;

Best Local Similarity 100.0%; Pred. No. 7e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
Db 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 E 61
Db 61 E 61

RESULT 7

Q69408 PRELIMINARY; PRT; 864 AA.
AC Q69408;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GB homolog.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FC126;
RX MEDLINE=94233711; PubMed=8178437;
RA Yoshida S., Lee L.F., Yanagida N., Nazerian K.;
RT "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
RT identification and expression by recombinant fowlpox viruses.";
RL Virology 200; 484-493 (1994).
DR EMBL; U01887; AAA19447.1; -;
DR InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.

DR Prodom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 864 AA; 98168 MW; 74060C71948F4B7B CRC64;

Query Match 56.4%; Score 184; DB 12; Length 864;
Best Local Similarity 57.4%; Pred. No. 5e-15;
Matches 35; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
Db 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 E 61
Db 61 E 61

RESULT 8

Q9DPO9 PRELIMINARY; PRT; 870 AA.
AC Q9DPO9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UL27 virion membrane glycoprotein B.
GN HVT035.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FC126;
RX MEDLINE=20578232; PubMed=11134310;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
RT "The genome of turkey herpesvirus.";
RL J. Virol. 75:971-978 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-FC126;
RX Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF291866; AAG45765.1; -;
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR Prodom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 870 AA; 98816 MW; F93D1C7036938FB1 CRC64;

Query Match 56.4%; Score 184; DB 12; Length 870;

Best Local Similarity 57.4%; Pred. No. 5.1e-15;
Matches 35; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 60
Db 7 MHYFRNCIFFLIVLYGTNSPSTQNTSVREVSSVOLSEESTFYLCPPVPGSTVIRL 66
QY 61 E 61
Db 67 E 67

RESULT 9

Q9E1G4 PRELIMINARY; PRT; 870 AA.
AC Q9E1G4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UL27 glycoprotein gb.
GN UL27.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=FC126;
RX  MEDLINE=21195611; PubMed=11297687;
RA  Kingham B.F., Zelnik V., Kopacek J., Majerclak V., Ney E.,
RA  Schmidt C.J.;
RT  "The genome of herpesvirus of turkeys: comparative analysis with
RT  Marek's disease viruses.";
RL  J. Gen. Virol. 82:1123-1135(2001).
DR  EMBL; AF282130; AAG30067.1; -
DR  InterPro; IPR000234; Glycoprot_B.
DR  Pfam; PF00606; Glycoprotein_B; 1.
DR  ProDom; PD000693; Glycoprot_B; 1.
SQ  SEQUENCE 870 AA; 98827 MW; CE02PB7E7F9EAA2 CRC64;

Query Match          56.4%; Score 184; DB 12; Length 870;
Best Local Similarity 57.4%; Pred. No. 5.1e-15;
Matches 35; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY  1 MHYFRNCIFFLVILYGTNSSPSTQNTVSREVVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB  7 MKYFNRSFLIFLPILSIATSEIKLPNVTAREIVSGIQLSEDETTFYVCPPPGVSTVIRL 66
QY  61 E 61
DB  67 E 67

RESULT 10
Q69406 Q69406 PRELIMINARY; PRT; 865 AA.
AC  Q69406;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  GB homolog.
OS  Turkey herpesvirus.
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Alphaherpesvirinae; Marek's disease-like viruses.
OX  NCBI_TaxID=10390;
RN  SEQUENCE FROM N.A.
RC  STRAIN=SB-1;
RX  MEDLINE=94233711; PubMed=8178437;
RA  Yoshida S., Lee L.F., Yanagida N., Nazarian K.;
RT  "The glycoprotein B genes of Marek's disease virus serotypes 2 and 3:
RT  identification and expression by recombinant fowlpox viruses.";
RL  Virology 200:484-493(1994).
DR  EMBL; U01886; AAA19445.1; -
DR  InterPro; IPR000234; Glycoprot_B.
DR  Pfam; PF00606; Glycoprotein_B; 1.
DR  ProDom; PD000693; Glycoprot_B; 1.
SQ  SEQUENCE 865 AA; 97236 MW; 04D4802FDC2E3948 CRC64;

Query Match          52.6%; Score 171.5; DB 12; Length 865;
Best Local Similarity 56.5%; Pred. No. 2.1e-13;
Matches 35; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY  1 MHYFRNCI-FFLIVILYGTNSSPSTQNTVSREVVSSVOLSEESTFYLCPPPGVSTVIR 59
DB  1 MNHFGICVPLFISVVFYFGRVARAQNVTISREAVSSVOLSEESFFLCPPPAVGTTVR 60
QY  60 LE 61
DB  61 LE 62

RESULT 11
Q9PW21 Q9PW21 PRELIMINARY; PRT; 865 AA.
AC  Q9PW21;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```

```

DT  01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE  Homolog of HSV-1 major DNA-binding protein (Glycoprotein B).
GN  ORF 35 OR UL27.
OS  Turkey herpesvirus,
OS  Marek's disease virus serotype 2 MDV2, and
OS  Gallid herpesvirus 3.
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Alphaherpesvirinae; Marek's disease-like viruses.
OX  NCBI_TaxID=10390, 36353, 35250;
RN  SEQUENCE FROM N.A.
RC  SPECIES=Turkey herpesvirus; STRAIN=HPRS24;
RA  Kato K., Jang H., Izumiya Y., Cai J., Tsushima Y., Miyazawa T.,
RA  Kai C., Mikami T.;
RT  "Identification and Transcriptional Analysis of the Marek's Disease
RT  Virus Serotype 2 Genes Homologous to the Glycoprotein B (UL27), the
RT  ICP18.5 (UL28) and the Major DNA-binding Protein (UL29) Genes of
RT  Herpes Simplex Virus Type 1.";
RL  J. Vet. Med. Sci. 0:0-0(1999).
RN  SEQUENCE FROM N.A.
RC  SPECIES=Gallid herpesvirus 1; STRAIN=HPRS24;
RA  Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
RA  Lee Y., Kai C., Takahashi E., Mikami T.;
RT  "The complete DNA sequence and transcription map of the unique long
RT  genome region of Marek's disease virus type 2.";
RL  Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN  SEQUENCE FROM N.A.
RC  SPECIES=Gallid herpesvirus 3; STRAIN=HPRS24;
RA  Izumiya Y., Jang H., Ono M., Mikami T.;
RL  Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN  SEQUENCE FROM N.A.
RC  SPECIES=Gallid herpesvirus 3; STRAIN=HPRS24;
RA  Izumiya Y., Jang H., Ono M., Mikami T.;
RT  "A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
RT  Strain HPRS24.";
RL  Curr. Top. Microbiol. Immunol. 0:0-0(2000).
DR  EMBL; AB024711; BAA83751.1; -
DR  EMBL; AB024414; BAA82923.1; -
DR  EMBL; AB049735; BAB16537.1; -
DR  InterPro; IPR000234; Glycoprot_B.
DR  Pfam; PF00606; Glycoprotein_B; 1.
DR  ProDom; PD000693; Glycoprot_B; 1.
KW  DNA-binding.
SQ  SEQUENCE 865 AA; 97309 MW; A20E66C70A475C35 CRC64;

Query Match          52.6%; Score 171.5; DB 12; Length 865;
Best Local Similarity 56.5%; Pred. No. 2.1e-13;
Matches 35; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY  1 MHYFRNCI-FFLIVILYGTNSSPSTQNTVSREVVSSVOLSEESTFYLCPPPGVSTVIR 59
DB  1 MNHFGICVPLFISVVFYFGRVARAQNVTISREAVSSVOLSEESFFLCPPPAVGTTVR 60
QY  60 LE 61
DB  61 LE 62

RESULT 12
Q91SD0 Q91SD0 PRELIMINARY; PRT; 879 AA.
AC  Q91SD0;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE  Glycoprotein B.
GN  GB.
OS  Canine herpesvirus.
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Alphaherpesvirinae.

```

```
OX NCBI_TaxID=37110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Australian;
RA Reubel G.H., Pekin J., Webb-Wagg K., Hardy C.M.;
RT "Nucleotide sequence of glycoprotein genes B, D, G, H and I, thymidine
RT kinase and protein kinase genes and gene homolog UL 24 of an
RT Australian isolate of canine herpesvirus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF361073; AAK51052.1; -
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 879 AA; 99584 MW; D37E2C586D7E59A1 CRC64;

Query Match 33.0%; Score 107.5; DB 12; Length 879;
Best Local Similarity 35.7%; Pred. No. 4.3e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY 9 IFFLVIL-----YGTNSPSTQNTVTSREVSSVOL-SEESTYLCPP 51
DB 8 IFFIITLICDPTTPESTINPLNHNLSLTPKPTSDDIRILRESQIESDDTSTFYMCPP 67

QY 52 PVGSTVIRLE 61
DB 68 PSGSLVIRLE 77

RESULT 13
QJ052 PRELIMINARY; PRT; 891 AA.
AC QJ052
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GB glycoprotein.
GN UL27
OS Simian herpes B virus (Cercopithecoid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E90-136;
RA Eberle R., Black D.H.;
DR EMBL; AF226637; AAF36698.1; -
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 891 AA; 98750 MW; 88A8C9A8A58F9284 CRC64;

Query Match 32.8%; Score 107; DB 12; Length 891;
Best Local Similarity 40.9%; Pred. No. 5e-05;
Matches 18; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 18 GNSPSTQNTVTSREVSSVOLSEESTYLCPPVPGSTVIRLE 61
DB 71 GTNSVEAGHATLRENLRDKALDGDATFYVCPPTGATVQFE 114

RESULT 14
Q66018 PRELIMINARY; PRT; 891 AA.
AC Q66018
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GB glycoprotein.
GN UL27
OS Simian herpes B virus (Cercopithecoid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
```

```
OX NCBI_TaxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2490;
RX MEDLINE=95251550; PubMed=7733825;
RA "Herpesvirus papio 2, an SAg-like alpha-herpesvirus of baboons.";
RL Arch. Virol. 140:529-545(1995).
DR EMBL; U14664; AAA85652.1; -
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 891 AA; 98624 MW; 8D61234C1A6330CB CRC64;

Query Match 32.8%; Score 107; DB 12; Length 891;
Best Local Similarity 40.9%; Pred. No. 5e-05;
Matches 18; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 18 GNSPSTQNTVTSREVSSVOLSEESTYLCPPVPGSTVIRLE 61
DB 72 GTNSVEAGHATLRENLRDKALDGDATFYVCPPTGATVQFE 115

RESULT 15
Q90050 PRELIMINARY; PRT; 948 AA.
AC Q90050
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Herpes simplex virus glycoprotein B (98) homolog.
OS Feline herpesvirus (Fellid herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93090104; PubMed=1333759;
RA Maeda K., Horimoto T., Norimine J., Kawaguchi Y., Tomonaga K.,
RA Nilkura M., Kai C., Takahashi E., Mikami T.;
RT "Identification and nucleotide sequence of a gene in feline
RT herpesvirus type 1 homologous to the herpes simplex virus gene
RT encoding the glycoprotein B.";
RL Arch. Virol. 127:387-397(1992).
DR EMBL; S49775; AAB24381.2; -
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR PROSITE; PS00018; EF_HAND; 1.
SQ SEQUENCE 948 AA; 106233 MW; 71C28FB00B091325 CRC64;

Query Match 32.5%; Score 106; DB 12; Length 948;
Best Local Similarity 23.5%; Pred. No. 7.3e-05;
Matches 28; Conservative 14; Mismatches 17; Indels 60; Gaps 3;

QY 3 YFRNCIF-----FLVLY----- 17
DB 22 YFQRCFFPSLLGIATGSRHNGSSGLTRLARYSVFIWLVLPVGPVPGSGSTSEQ 81

QY 18 -----GTNSPSTQNTVTSREVSSVOLSEESTYLCPPVPGSTVIRLE 61
DB 82 PRTVATPEVGTPPKPTTDDTMDSDMRALRASQIEANGSPFFYMCPPPGSTVVRLE 140

Search completed: October 8, 2003, 17:07:01
Job time : 7.42292 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:35 ; Search time 1.13439 Seconds
(without alignments)
2611.703 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFFLIVLYGTN.....STFYLCPPVGVSTVIRLEFG 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	314	96.3	865	1	VGLB_HSVMD
2	100	30.7	868	1	VGLB_VZVD
3	96	29.4	885	1	VGLB_HSV2S
4	95	29.1	903	1	VGLB_HSV1F
5	95	29.1	904	1	VGLB_HSV1I
6	95	29.1	904	1	VGLB_HSV1K
7	95	29.1	904	1	VGLB_HSV1P
8	91	27.9	979	1	VGLB_HSV1L
9	91	27.9	980	1	VGLB_HSV1E
10	91	27.9	980	1	VGLB_HSV1A
11	91	27.9	980	1	VGLB_HSV1B
12	87.5	26.8	933	1	VGLB_HSV1C
13	87	26.7	917	1	VGLB_HSV1D
14	86	26.4	919	1	VGLB_HSV1E
15	85	26.1	904	1	VGLB_HSV1F
16	85	26.1	904	1	VGLB_HSV1G
17	84	25.8	920	1	VGLB_HSV1H
18	84	25.8	920	1	VGLB_HSV1I
19	82.5	25.3	928	1	VGLB_HSV1J
20	82.5	25.3	932	1	VGLB_HSV1K
21	63	19.3	649	1	GUND_CLOTH
22	60.5	18.6	561	1	EST4_RAT
23	59	18.1	1436	1	DPO3_STAAW
24	59	18.1	1438	1	DPO3_STAAW
25	58	17.8	1223	1	K125_CAEEL
26	57.5	17.6	281	1	UPK_STRPN
27	57	17.3	1578	1	NX3A_RAT
28	56.5	17.3	250	1	PHON_SALTY
29	56	17.2	251	1	ISTB_BAGST
30	56	17.2	378	1	CKD6_MOUSE
31	56	17.2	655	1	NEC3_MOUSE
32	55.5	17.0	248	1	GVFJ_ANAFL
33	55.5	17.0	456	1	NX3B_BOVIN

34 55.5 17.0 563 1 HEMA_IAMAA
35 55.5 17.0 637 1 FTSH_BACSU
36 55 16.9 104 1 YDTH_SCHPO
37 55 16.9 302 1 SPRC_RAT
38 55 16.9 382 1 CRPC_CHLTE
39 55 16.9 430 1 WR14_ARATH
40 55 16.9 968 1 PKD2_HUMAN
41 55 16.9 1567 1 ICEN_XANCT
42 54.5 16.7 568 1 HEMA_IAMAB
43 54 16.6 470 1 YEF7_YEAST
44 54 16.6 549 1 TEGU_HCMV
45 54 16.6 654 1 HEMA_INCCA

ALIGNMENTS

RESULT 1
VGLB_HSVMD
ID VGLB_HSVMD STANDARD; PRT; 865 AA.
AC P18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID:33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:89293086; PubMed:2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;
RT "Nucleotide sequence and characterization of the Marek's disease
virus homologue of glycoprotein B of herpes simplex virus.";
RL J. Gen. Virol. 70:1789-1804(1989).
CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13713; BAA02866.1; -
CC InterPro; IPR000234; Glycoprot_B.
CC Pfam; PF00606; Glycoprotein_B; 1.
CC ProDom; PD000693; Glycoprot_B; 1.
KW Signal; Glycoprotein; Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 865 GLYCOPROTEIN B.
FT DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 683 700 POTENTIAL.
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 732 752 POTENTIAL.
FT DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 98091 MW; B30E93CIAC65C6C3 CRC64;

Query Match 96.3%; Score 314; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 9.3e-31;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHYFRNCIFFLIVLYGTN.....STFYLCPPVGVSTVIRLEFG 60
|||||

```

Db      1 MHYFARNCIFFLIVILVIGTNSPSTQNTSREVVSQVLSSEESTFYLCPPVGVSTVIRL 60
Qy      61 E 61
Db      61 E 61

RESULT 2
VGLB_VZVD
ID  VGLB_VZVD      STANDARD;      PRT;      868 AA.
AC  P09257;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-MAR-1989 (Rel. 10, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Glycoprotein B precursor (Glycoprotein II).
GN  31.
OS  varicella-zoster virus (strain Dumas) (VZV).
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=10338;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86306657; PubMed=3018124;
RA  Davison A.J., Scott J.E.;
RT  "The complete DNA sequence of varicella-zoster virus.";
RL  J. Gen. Virol. 67:1759-1816(1986).
CC  -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; X04370; CAA27914.1; -.
CC  PIR; E27214; VGBE31.
DR  InterPro; IPR000234; Glycoprot_B.
DR  Pfam; PF00606; Glycoprotein_B; 1.
DR  ProDom; PD000693; Glycoprot_B; 1.
DR  Glycoprotein; Transmembrane; Signal.
FT  SIGNAL      1      2
FT  CHAIN       ?      868      GLYCOPROTEIN B.
FT  DOMAIN      ?      679      EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM    680      695      POTENTIAL.
FT  TRANSMEM    701      720      POTENTIAL.
FT  TRANSMEM    724      744      POTENTIAL.
FT  DOMAIN      745      868      CYTOPLASMIC (POTENTIAL).
FT  CARBOHYD     84      84      N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD    194      194      N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD    372      372      N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD    440      440      N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD    557      557      N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD    623      623      N-LINKED (GLCNAC. .) (POTENTIAL).
SQ  SEQUENCE      868 AA; 98066 MW; 58B8E1346FAFF902 CRC64;

Query Match      30.7%; Score 100; DB 1; Length 868;
Best Local Similarity 54.8%; Pred. No. 0.00023;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy      31 REVVSQVLSSEESTFYLCPPVGVSTVIRLE 61
Db      41 REAIHKSQDAETKPTFYVCPPTGVSTVIRLE 71

RESULT 3
VGLB_HSV2S
ID  VGLB_HSV2S      STANDARD;      PRT;      885 AA.
AC  P24954;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-MAR-1992 (Rel. 21, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)

```

```

RA Pellett P.E., Kousoulas K.G., Pereira L., Roizman B.;
RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
RT primary sequence and predicted protein structure of the wild type and
RT of monoclonal antibody-resistant mutants.";
RL J. Virol. 53:243-253(1985).
RN [2]
RP SEQUENCE OF 1-176 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hammerichmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -|- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -|- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -|- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14164; AAA45776.1; -.
DR EMBL: M21633; AAA45788.1; -.
DR PIR: A03750; VGBEB1.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 29
FT CHAIN 30 GLYCOPROTEIN B.
FT DOMAIN 31 903 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 730 745 POTENTIAL.
FT TRANSMEM 751 770 POTENTIAL.
FT TRANSMEM 774 794 POTENTIAL.
FT DOMAIN 795 903 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 903 AA; 100104 MW; 739DC47813DB35E8 CRC64;

Query Match 29.1%; Score 95; DB 1; Length 903;
Best Local Similarity 34.1%; Pred. No. 0.001;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 18 GTNSSPSTQNTVTSREVSVSVQSEESTFYLCPPPVGVSTVIRLE 61
DB 84 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVVGQFE 127

RESULT 5
VGLB_HSV11
ID VGLB_HSV11 STANDARD; PRT; 904 AA.
AC P10211;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
CC -|- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -|- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -|- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14112; CAA32320.1; -.
DR PIR: I30084; VGBEW7.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 904 GLYCOPROTEIN B.
FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 731 746 POTENTIAL.
FT TRANSMEM 752 771 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT DOMAIN 796 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100292 MW; 2C14E8B1284C1F3A CRC64;

Query Match 29.1%; Score 95; DB 1; Length 904;
Best Local Similarity 34.1%; Pred. No. 0.001;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 18 GTNSSPSTQNTVTSREVSVSVQSEESTFYLCPPPVGVSTVIRLE 61
DB 85 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVVGQFE 128

RESULT 6
VGLB_HSV1K
ID VGLB_HSV1K STANDARD; PRT; 904 AA.
AC P06437;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84174058; PubMed=6324454;
RA Bzik D.J., Fox B.A., Deluca N.A., Person S.;
RT "Nucleotide sequence specifying the glycoprotein gene, gb, of herpes
RT simplex virus type 1.";
RL Virology 133:301-314(1984).

```

```

RN REVISIONS. SEQUENCE FROM N.A.
RX MEDLINE-87071654; PubMed-3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gp glycoprotein gene of HSV-2 and
RL comparison with the corresponding gene of HSV-1.";
RN Virology 155:322-333(1986).
RN REVISIONS.
RA Pederson N.E.;
RL Submitted (APR-1987) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K01760; AAA45774.1;
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 904 GLYCOPROTEIN B.
FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 731 746 POTENTIAL.
FT TRANSMEM 752 771 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT DOMAIN 796 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100368 MW; B97B7F8DE5FBA299 CRC64;

Query Match 29.1%; Score 95; DB 1; Length 904;
Best Local Similarity 34.1%; Pred. No. 0.001;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFVLCPPPVGVSTVIRLE 61
| : : : | : : : : : : : : : : : : : : : : : : : : : : :
Db 85 GDNATVAAGHATLREHLRDKAENTDANFVCPPTGATVVOQE 128

RESULT 7
VGLB_HSV1P
ID VGLB_HSV1P STANDARD; PRT; 904 AA.
AC P08665;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain Patton).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10308;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87112925; PubMed-3027364;
RA Stuve L.L., Brown-Shimer S., Pachl C., Najarian R., Dina D.,
RA Burke L.L.;

```

```

RT "Structure and expression of the herpes simplex virus type 2
RL glycoprotein gB gene.";
RN J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K03541; AAA5778.1;
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 904 GLYCOPROTEIN B.
FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 731 746 POTENTIAL.
FT TRANSMEM 752 772 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT DOMAIN 796 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100115 MW; 7825E1DC830A626F CRC64;

Query Match 29.1%; Score 95; DB 1; Length 904;
Best Local Similarity 34.1%; Pred. No. 0.001;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFVLCPPPVGVSTVIRLE 61
| : : : | : : : : : : : : : : : : : : : : : : : : : : :
Db 85 GDNATVAAGHATLREHLRDKAENTDANFVCPPTGATVVOQE 128

RESULT 8
VGLB_HSV1P
ID VGLB_HSV1P STANDARD; PRT; 979 AA.
AC P25218;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10330;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90236317; PubMed-1692002;
RA Guo P.;
RT "Characterization of the gene and an antigenic determinant of equine
RL herpesvirus type-1 glycoprotein 14 with homology to gB-equivalent
RL glycoproteins of other herpesviruses.";
RN Gene 87:249-255(1990).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M34861; AAA46086.1; -
 DR PIR; JH0109; JH0109
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 979
 FT DOMAIN 87 851
 FT TRANSMEM 852 869
 FT DOMAIN 870 979
 FT CARBOHYD 165 165
 FT CARBOHYD 275 275
 FT CARBOHYD 380 380
 FT CARBOHYD 423 423
 FT CARBOHYD 497 497
 FT CARBOHYD 514 514
 FT CARBOHYD 515 515
 FT CARBOHYD 560 560
 FT CARBOHYD 727 727
 FT CARBOHYD 749 749
 SQ SEQUENCE 979 AA; 110331 MW; 9A19866B791C5B36 CRC64;

Query Match 27.9%; Score 91; DB 1; Length 979;
 Best Local Similarity 28.7%; Pred. No. 0.0035;
 Matches 25; Conservative 7; Mismatches 25; Indels 30; Gaps 2;

QY 5 RRCIFFLVLVYG-----TNSPSTQNVTS-----REV 34
 DB 66 RVGTIVLACLLFGSCVVRVPTPTPTSTPTSMTHSHGTVDPDLLTPTDPLRLAV 125
 QY 35 SSVOLSEESTFYLCPPPGVSTVIRLE 61
 DB 126 RESGILAEADGDFYTCPPPTGTVVRIE 152

RESULT 9
 VGLB_HSV1
 ID VGLB_HSV1 STANDARD; PRT; 980 AA.
 AC P18050;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33.
 OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10327;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89279217; PubMed=2543744;
 RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
 RA Woodworth L.M.;
 RT Identification and nucleotide sequence of a gene in equine
 RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
 RT major envelope glycoprotein gB.;
 RL J. Gen. Virol. 70:383-394(1989).
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D00401; BAA00304.1; ALT_SEQ.
 DR PIR; A31241; VGBE2H.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 980
 FT DOMAIN 87 852
 FT TRANSMEM 853 870
 FT DOMAIN 871 980
 FT CARBOHYD 165 165
 FT CARBOHYD 275 275
 FT CARBOHYD 380 380
 FT CARBOHYD 423 423
 FT CARBOHYD 497 497
 FT CARBOHYD 514 514
 FT CARBOHYD 515 515
 FT CARBOHYD 560 560
 FT CARBOHYD 727 727
 FT CARBOHYD 749 749
 SQ SEQUENCE 980 AA; 109931 MW; 94B69AF882FB6AC2 CRC64;

Query Match 27.9%; Score 91; DB 1; Length 980;
 Best Local Similarity 28.7%; Pred. No. 0.0035;
 Matches 25; Conservative 7; Mismatches 25; Indels 30; Gaps 2;

QY 5 RRCIFFLVLVYG-----TNSPSTQNVTS-----REV 34
 DB 66 RVGTIVLACLLFGSCVVRVPTPTPTSTPTSMTHSHGTVDPDLLTPTDPLRLAV 125
 QY 35 SSVOLSEESTFYLCPPPGVSTVIRLE 61
 DB 126 RESGILAEADGDFYTCPPPTGTVVRIE 152

RESULT 10
 VGLB_HSVEA
 ID VGLB_HSVEA STANDARD; PRT; 980 AA.
 AC P18551;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33.
 OS Equine herpesvirus type 1 (strain AB1) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10328;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bonass W.A., Elton D.M., Stocks J.M., Killington R.A.,
 RA Meredith D.M., Halliburton I.W.;
 RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

[illegible]

```

QY      35  SSVOLSEESTFYLCPPPGVGSTVIRLE 61
      : 1:  || |||| |||||:|:|
Db      126  RESGILAEQDGYTCPPPTGTVVRIE 152

RESULT 12
VGLB_HSVAI
ID      VGLB_HSVAI      STANDARD;      PRT;      933  AA.
AC      Q04463;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Glycoprotein B precursor.
GN      UL27.
OS      Herpesvirus ateles (type 1 / strain Lennette).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Alphaherpesvirinae.
OX      NCBI_TaxID=35243;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9328440; PubMed=8385913;
RA      Eberlin R., Black D.;
RT      "sequence analysis of herpes simplex virus gb gene homologs of two
RT      platyrrhine monkey alpha-herpesviruses.";
RL      Arch. Virol. 129:167-182(1993).
CC      -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC      REQUIRED FOR VIRAL GROWTH.
CC      -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC      -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-30 IS THE INITIATOR.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announcement
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: M95785; AAA43839.1; -.
CC      InterPro: IPR000234; Glycoprot_B.
CC      Pfam: PF00606; Glycoprotein_B.1.
CC      ProDom: PD000693; Glycoprot_B.1.
KW      Glycoprotein; Transmembrane; Signal.
FT      SIGNAL 1 31 POTENTIAL.
FT      CHAIN 32 933 GLYCOPROTEIN B.
FT      DOMAIN 32 753 EXTRACELLULAR (POTENTIAL)

```

```

FT TRANSMEM 754 769 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT TRANSMEM 798 818 POTENTIAL.
FT DOMAIN 819 933 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 933 AA; 101492 MW; 13A6CF88B5A18DF CRC64;

Query Match 26.8%; Score 87.5; DB 1; Length 933;
Best Local Similarity 40.9%; Pred. No. 0.0089;
Matches 18; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

Qy 18 GTNSSPSTQNTSREVSSVQLSEESTFYLCPPVPGSTVIRLE 61
|||: | | | | | | | | | | | | | | | | | | | |
Db 108 GTRSAARRQ---LRESLRRIQAEYAASAFVVCPPPTGATVVQFE 148

RESULT 13
VGLB_HSVB2
ID VGLB_HSVB2 STANDARD; PRT; 917 AA.
AC P12641;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B-1 precursor.
OS Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88306231; PubMed=2841793;
RA Hammerschmidt W., Contraths F., Mankertz J., Pauli G., Ludwig H.,
RA Buhk H.-J.;
RT "Conservation of a gene cluster including glycoprotein B in bovine
herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";
RL Virology 165:388-405(1988).
RN [2]
RP SEQUENCE OF 1-200 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hammerschmidt W., Contraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
proteins of bovine herpesvirus type 2 (BHV-2) and herpes
virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -1- FUNCTION: GBL IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE
PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A
SYNCYTIAL PHENOTYPE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).

```

```

FT TRANSMEM 770 786 POTENTIAL.
FT TRANSMEM 795 811 POTENTIAL.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 917 AA; 101882 MW; 1B96CBF50DB4D3F3 CRC64;

Query Match 26.7%; Score 87; DB 1; Length 917;
Best Local Similarity 38.1%; Pred. No. 0.01;
Matches 16; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

Qy 20 NSSPSTQNTSREVSSVQLSEESTFYLCPPVPGSTVIRLE 61
|||: | | | | | | | | | | | | | | | | | | | |
Db 110 NASEPADPAELRADRLGLKSSDDPNFYVCPPTGATVVRL 151

RESULT 14
VGLB_HSVE4
ID VGLB_HSVE4 STANDARD; PRT; 919 AA.
AC P17472;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89123704; PubMed=2915378;
RA Riggiio M.P., Cullinane A.A., Onions D.E.;
RT "Identification and nucleotide sequence of the glycoprotein gb gene
of equine herpesvirus 4.";
RL J. Virol. 63:1123-1133(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).

```

```
FT CARBOHYD 688      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 919 AA; 103710 MW; C77E48F26C37BC7B CRC64;

Query Match      26.4%; Score 86; DB 1; Length 919;
Best Local Similarity 35.0%; Pred. No. 0.013;
Matches 21; Conservative 7; Mismatches 14; Indels 18; Gaps 2;

QY 20 NSSPST-----QNVTSREVSSVOLSEES-----TFYLCPPPVGTVIRLE 61
DB 34 SSQSPSTPASTQSAKTVDTLLPTETPDPLRLAVRESGILAEADGDFYTCPPPTGTSTVVRIE 93

Search completed: October 8, 2003, 17:03:36
Job time : 2.13439 secs

RESULT 15
VGLB_HSV23
ID VGLB_HSV23 STANDARD; PRT; 904 AA.
AC P06763;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27 OR GB2.
OS Herpes simplex virus (type 2 / strain 333).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Alphaherpesvirinae; Simplexvirus.
RN NCBI_TaxID=10313;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87112925; PubMed=3027364;
RA Stuve L.L., Brown-Shimer S., Pachl C., Najarian R., Dina D.,
RA Burke R.L.;
RT "Structure and expression of the herpes simplex virus type 2
RT glycoprotein gb gene."
RL J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15118; AAA45837.1; -.
DR PIR; A26790; VGBEB2.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 904 GLYCOPROTEIN B.
FT DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 728 743 POTENTIAL.
FT TRANSMEM 749 768 POTENTIAL.
FT TRANSMEM 772 792 POTENTIAL.
FT DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100186 MW; A8B36F74FDBC539 CRC64;

Query Match      26.1%; Score 85; DB 1; Length 904;
Best Local Similarity 31.0%; Pred. No. 0.017;
Matches 13; Conservative 14; Mismatches 15; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:00 ; Search time 2.1581 Seconds
(without alignments)
2807.386 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFFLIVLYGTN.....STFYLCPPVGVSTVIRLEFG 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	314	96.3	865	1	VGBERB		glycoprotein B pre
2	106	32.5	948	2	A56602		glycoprotein B hom
3	100	30.7	868	1	VGBE31		glycoprotein B - h
4	99.5	30.5	943	2	B48474		glycoprotein B - f
5	96	29.4	885	1	VGBE5A		glycoprotein B pre
6	95	29.1	903	1	VGBE51		glycoprotein B pre
7	95	29.1	903	1	VGBEK1		glycoprotein B pre
8	95	29.1	904	1	VGBEW7		glycoprotein B pre
9	91	27.9	979	2	JH0109		glycoprotein 14 pr
10	91	27.9	980	1	VGBE2H		glycoprotein B pre
11	91	27.9	980	1	VGBEK2		glycoprotein B pre
12	87.5	26.8	933	1	B48349		glycoprotein B pre
13	87	26.7	917	1	VGBE5H		glycoprotein B pre
14	86	26.4	919	1	VGBEQH		glycoprotein B pre
15	86	26.4	975	2	T42576		probable envelope
16	85	26.1	904	1	VGBEK2		glycoprotein B pre
17	85	26.1	904	1	VGBE52		glycoprotein B pre
18	84	25.8	913	1	VGBE5S		glycoprotein gII p
19	82.5	25.3	928	1	VGBE5G		glycoprotein gI pr
20	82.5	25.3	932	1	VGBE5C		glycoprotein gI pr
21	82	25.2	920	1	D48349		glycoprotein B pre
22	65	19.9	392	2	F96707		hypothetical prote
23	64.5	19.8	497	2	T19077		hypothetical prote
24	63.5	19.5	207	2	A12505		hypothetical prote
25	63	19.3	649	1	C2CLDM		cellulase (EC 3.2.
26	62	19.0	312	2	T01250		hypothetical prote
27	60.5	18.6	561	2	S62788		carboxylesterase (
28	60.5	18.6	643	2	H88712		protein CL7H12.1 [
29	60	18.4	343	2	T09565		peroxidase (EC 1.1

30 59 18.1 954 2 G86312
31 59 18.1 1415 2 S52267
32 59 18.1 1438 2 C89900
33 58.5 17.9 78 2 JH0344
34 58.5 17.9 296 2 JC7268
35 58.5 17.9 1797 2 T21889
36 58.5 17.9 1805 2 T21888
37 58 17.8 64 2 C70235
38 58 17.8 209 2 D71069
39 58 17.8 411 2 S11961
40 58 17.8 1223 2 T15316
41 58 17.8 1471 2 B48218
42 57.5 17.6 281 2 H95052
43 57.5 17.6 281 2 E97923
44 57.5 17.6 284 2 T22671
45 57.5 17.6 336 2 T22680

ALIGNMENTS

RESULT 1

VGBERB

glycoprotein B precursor - Marek's disease virus (strain RBLB)

C:Species: Marek's disease virus

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C:Accession: A32402; B32402

R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog

A:Reference number: A32402; MUID:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 <ROS>

A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BA02866.1; PID:g221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271;304-330 <ROS2>

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-865/Product: glycoprotein B #status predicted <GPB>

F:709-728/Domain: transmembrane #status predicted <TN1>

F:732-752/Domain: transmembrane #status predicted <TN2>

F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 96.3% Score 314; DB 1; Length 865;

Best Local Similarity 100.0%; Pred. No. 9.4e-29;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHYFRNCIFFLIVLYGTNSPQTQNTSREWSSVOLSEESTFYLCPPVGVSTVIRL 60

Db 1 MHYFRNCIFFLIVLYGTNSPQTQNTSREWSSVOLSEESTFYLCPPVGVSTVIRL 60

Qy 61 E 61

Db 61 E 61

RESULT 2

A56602

glycoprotein B homolog precursor - feline herpesvirus 1

C:Species: feline herpesvirus 1

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999

C:Accession: A56602

R:Maeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Nukura, M.; Ka

Arch. Virol. 127, 387-397, 1992

A:Title: Identification and nucleotide sequence of a gene in feline herpesvirus type

A:Reference number: A56602; MUID:93090104; PMID:1333759

A:Accession: A56602

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-948 <MAE>

A;Cross-references: CB:S49775; NID:g261094; PIDN:AAB24381.l.; PID:g261095
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:120002, NCBIP:120003)
C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein

Query Match 32.5%; Score 106; DB 2; Length 948;
Best Local Similarity 23.5%; Pred. No. 0.00028;
Matches 28; Conservative 14; Mismatches 17; Indels 60; Gaps 3;

QY 3 YFRNRCLF-----FLIVLY----- 17
|||: ||

Db 22 YRQRCEFPPLGILGAATGRSHNGSSGLTRIAKYVSFIWVLEVLVGPRPVEGGSGSTSEQ 81
||||:

QY 18 -----GNSSPSNTVTREYVVSVQLSEEE----STFYLCPPPVGVGSTVIRLE 61
|| | : : : ||

Db 82 PRRTVATPEVGGTPPKPTDPTOMSDMRALRASQIEANGPFSTFMCPFGSSTVVRLE 140
|| | : : : |||:

RESULT 3
VGBE31
glycoprotein B - human herpesvirus 3
N;Alternate names: glycoprotein II
C;Species: human herpesvirus 3, varicella-zoster virus
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C;Accession: E27214
R;Davison, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A;Title: The complete DNA sequence of varicella-zoster virus.
A;Reference number: A27345; UID:86306657; PMID:3018124
A;Accession: E27214
A;Molecule type: DNA
A;Residues: 1-868 <DAV>
C;Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27914.1; PID:g60020
C;Genetics:
A;Gene: 31
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;697-713/Domain: transmembrane #status predicted <TM1>
F;726-742/Domain: transmembrane #status predicted <TM2>
F;84,194,372,416,440,494,557,623,781,836,864/Binding site: carbohydrate (Asn) (covalent)

Query Match 30.7%; Score 100; DB 1; Length 868;
Best Local Similarity 54.8%; Pred. No. 0.0013;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 31 REYVSSVOLSEEESTFYLCPPVPVGVSTVIRLE 61
|| | : : : |||:

Db 41 RRAHKSQADETFTPYVCPPPGTGSITVRL 71
|| | : : : |||:

RESULT 4
B48474
glycoprotein B - feline herpesvirus 1 (fragment)
C;Species: feline herpesvirus 1
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B48474
R;Spatz, S.J.; Maes, R.K.
Virology 197, 125-136, 1993
A;Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B and A
A;Reference number: A48474; UID:94025559; PMID:8212548
A;Contents: C-27
A;Accession: B48474
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-943 <SPA>
A;Cross-references: GB:S66371; NID:g435792; PIDN:AAB28559.1; PID:g435794
A;Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBIP:138807)
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein

Query Match 30.5%; Score 99.5; DB 2; Length 943;
Best Local Similarity 24.6%; Pred. No. 0.0016;

```
Db      84 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVVQFE 127
      | | : : : | | | : : : | | : | | | | | : | | : |
RESULT 7
VGBEK1
glycoprotein B precursor - human herpesvirus 1 (strain KOS)
C:Species: human herpesvirus 1
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03751
Virology 133, 301-314, 1984
R:Bzik, D.J.; Fox, B.A.; Deluca, N.A.; Person, S.
A:Title: Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus 1
A:Reference number: A03751; MUID:84174058; PMID:6324454
A:Accession: A03751
A:Molecule type: DNA
A:Residues: 1-903 <BZI>
A:Cross-references: GB:K01760; NID:g330082; PIDN:AAA45774.1; PID:g330083
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-903/Product: glycoprotein B #status predicted <MAT>
F:726-746/Domain: transmembrane #status predicted <TM1>
F:751-771/Domain: transmembrane #status predicted <TM2>
F:774-794/Domain: transmembrane #status predicted <TM3>
F:76,86,140,254,397,429,477,488,673,818,887/Binding site: carbohydrate (Asn) (covalent)
F:115-572,132-528,206-270,363-411,595-632/Disulfide bonds: #status predicted

Query Match      29.1%; Score 95; DB 1; Length 903;
Best Local Similarity 34.1%; Pred. No. 0.0053;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Qy      18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPPVGSTVIRLE 61
      | | : : : | | | : : : | | : | | | | | : | | : |
Db      84 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVVQFE 127

RESULT 8
VGBEW7
glycoprotein B precursor - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: I30084
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perry, J. Gen. Virol. 69, 1531-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes simplex virus 1
A:Reference number: A30083; MUID:88274327; PMID:2839594
A:Accession: I30084
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-904 <MCG>
A:Cross-references: GB:X14112; GB:D00317; GB:D00374; GB:S40593; NID:g1944536; PIDN:CAA32
C:Genetics:
A:Gene: UL27
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-904/Product: glycoprotein B #status predicted <MAT>
F:727-746/Domain: transmembrane #status predicted <TM1>
F:752-771/Domain: transmembrane #status predicted <TM2>
F:774-795/Domain: transmembrane #status predicted <TM3>
F:87,141,255,398,430,478,489,674,819,888/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:116-573,133-529,207-271,364-412,596-633/Disulfide bonds: #status predicted

Query Match      29.1%; Score 95; DB 1; Length 904;
Best Local Similarity 34.1%; Pred. No. 0.0053;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Qy      18 GTNSSPSTQNTVTSREVSVSSVOLSEESTFYLCPPPVGSTVIRLE 61
      | | : : : | | | : : : | | : | | | | | : | | : |
Db      85 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVVQFE 128
```

```
RESULT 9
JH0109
glycoprotein 14 precursor - equine herpesvirus 1
C:Species: equine herpesvirus 1
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 26-Aug-1999
C:Accession: JH0109
R:Guo, P.
Gene 87, 249-255, 1990
A:Title: Characterization of the gene and an antigenic determinant of equine herpesvirus 1
A:Reference number: JH0109; MUID:90236317; PMID:1692002
A:Accession: JH0109
A:Molecule type: DNA
A:Residues: 1-979 <GUO>
A:Cross-references: GB:M34861; NID:g330900; PIDN:AAA46086.1; PID:g330901
A:Note: Glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies, C:Genetics:
A:Gene: gp14
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-86/Domain: signal sequence #status predicted <SIG>
F:87-979/Product: glycoprotein 14 #status predicted <MAT>
F:831-871/Region: hydrophobic
F:165,275,380,423,497,514,560,727,749,951,970/Binding site: carbohydrate (Asn) (covalent)

Query Match      27.9%; Score 91; DB 2; Length 979;
Best Local Similarity 28.7%; Pred. No. 0.017;
Matches 25; Conservative 7; Mismatches 25; Indels 30; Gaps 2;

Qy      5 RRNCIFFLIVLYG-----TNSSPSTQNTVS-----REVW 34
      | | : | : | | | | | | | | | | | | | | | |
Db      66 RVGTIVLACLLFGSCVVRAVPTTSPPTSTPTSMTHSHGCTVDPDLLPTETPDPLRLAV 125

Qy      35 SSVOLSEESTFYLCPPPVGSTVIRLE 61
      : | : | | | | | | | | | | | | | | | |
Db      126 RESGILAEKDFYTCPPPTGTVVRIE 152

RESULT 10
VGBEZH
glycoprotein B precursor - equine herpesvirus 1 (isolate HVS 25A)
C:Species: equine herpesvirus 1
C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-Jan-1996
C:Accession: A31241; J00008
R:Whalley, J.M.; Robertson, G.R.; Scott, N.A.; Hudson, G.C.; Bell, C.W.; Woodworth, L. J. Gen. Virol. 70, 383-394, 1989
A:Title: Identification and nucleotide sequence of a gene in equine herpesvirus 1
A:Reference number: A31241; MUID:89279217; PMID:2543744
A:Accession: A31241
A:Molecule type: DNA
A:Residues: 1-980 <WHA>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-85/Domain: signal sequence #status predicted <SIG>
F:70-78/Domain: transmembrane #status predicted <TM1>
F:86-980/Product: glycoprotein B #status predicted <MAT>
F:853-868/Domain: transmembrane #status predicted <TN2>
F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (C

Query Match      27.9%; Score 91; DB 1; Length 980;
Best Local Similarity 28.7%; Pred. No. 0.017;
Matches 25; Conservative 7; Mismatches 25; Indels 30; Gaps 2;

Qy      5 RRNCIFFLIVLYG-----TNSSPSTQNTVS-----REVW 34
      | | : | : | | | | | | | | | | | | | | | |
Db      66 RVGTIVLACLLFGSCVVRAVPTTSPPTSTPTSMTHSHGCTVDPDLLPTETPDPLRLAV 125

Qy      35 SSVOLSEESTFYLCPPPVGSTVIRLE 61
      : | : | | | | | | | | | | | | | | | |
Db      126 RESGILAEKDFYTCPPPTGTVVRIE 152

RESULT 11
VGBEC6
```

glycoprotein B precursor - equine herpesvirus 1 (strain Ab4p)
C:Species: equine herpesvirus 1
A:Note: host Equus caballus (domestic horse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: G36798
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Submitted to GenBank, March 1992
A:Description: The DNA sequence of equine herpesvirus-1.
A:Reference number: A36805
A:Accession: G36798
A:Molecule type: DNA
A:Residues: 1-980 <TEL>
A:Cross-references: GB:M86664; NID:G330791; PIDN:AA802468.1; PID:G330825
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virol. 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295366; PMID:1318606
A:Contents: annotation: possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 33
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-85/Domain: signal sequence #status predicted <SIG>
F:67-78/Domain: transmembrane #status predicted <TM1>
F:86-980/Product: glycoprotein B #status predicted <MAT>
F:850-868/Domain: transmembrane #status predicted <TM2>
F:165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (covalent)
Query Match 27.9%; Score 91; DB 1; Length 980;
Best Local Similarity 28.7%; Pred. No. 0.017;
Matches 25; Conservative 7; Mismatches 25; Indels 25; Gaps 2;
QY 5 RNCIFFLIVILYG-----TNSPSTQNTVS-----REVV 34
DB 66 RVGTIVLACLLFGLGSCVNAVPTPTSTSMTHSHGVDPDTPDPLRLAV 125
QY 35 SSVQLSEESTFYLCPPVPGSVIRLE 61
DB 126 RESGILAEEDGDFYTCPPPTGTSVIRLE 152
RESULT 12
B48349
glycoprotein B precursor - ateline herpesvirus 1 (strain Lennette)
C:Species: ateline herpesvirus 1
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
C:Accession: B48349
R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A:Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine m
A:Reference number: A48349; MUID:93228440; PMID:838913
A:Accession: B48349
A:Molecule type: DNA
A:Residues: 1-933 <EBE>
A:Note: sequence extracted from NCBI backbone (NCBIN:129063, NCBI:129065)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-933/Product: glycoprotein B #status predicted <MAT>
F:775-794/Domain: transmembrane #status predicted <TM1>
F:801-818/Domain: transmembrane #status predicted <TM2>
F:107,161,418,450,697,747/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 26.8%; Score 87.5; DB 1; Length 933;
Best Local Similarity 40.9%; Pred. No. 0.042;
Matches 18; Conservative 7; Mismatches 16; Indels 3; Gaps 1;
QY 18 GTNSPSTQNTVSREVSSVQLSEESTFYLCPPVPGSVIRLE 61
DB 108 GTRSARHQ---LRSLRRIOAEYAAAFYCPPTGATVQVFE 148

glycoprotein B precursor - bovine herpesvirus 2 (strain BMV)
C:Species: bovine herpesvirus 2
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: C29242
R:Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.
Virol. 165, 388-405, 1988
A:Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus
A:Reference number: A94381; MUID:88306231; PMID:2841793
A:Accession: C29242
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-917 <HAM>
A:Cross-references: GB:M21628; NID:G330752; PIDN:AAA46053.1; PID:G330753
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-917/Product: glycoprotein B #status predicted <GPB>
F:578-594/Domain: transmembrane #status predicted <TM1>
F:770-786/Domain: transmembrane #status predicted <TM2>
F:795-811/Domain: transmembrane #status predicted <TM3>
F:48,110,164,278,421,453,505,564,692/Binding site: carbohydrate (Asn) (covalent) #sta
Query Match 26.7%; Score 87; DB 1; Length 917;
Best Local Similarity 38.1%; Pred. No. 0.047;
Matches 16; Conservative 9; Mismatches 17; Indels 0; Gaps 0;
QY 20 NSSPSTQNTVSREVSSVQLSEESTFYLCPPVPGSVIRLE 61
DB 110 NASEPADPAELRADLRLGLKSSDDPNFYVCPPPTGATVVRLE 151
RESULT 14
VBEBQH
glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C:Species: equine herpesvirus 4
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A31880
R:Riggio, M.P.; Cullinane, A.A.; Onions, D.E.
J. Virol. 63, 1123-1133, 1989
A:Title: Identification and nucleotide sequence of the glycoprotein gB gene of equine
A:Reference number: A31880; MUID:89125704; PMID:2915378
A:Accession: A31880
A:Molecule type: DNA
A:Residues: 1-919 <RIG>
A:Cross-references: GB:M26171; NID:G341446; PIDN:AAA46106.1; PID:G514920
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-919/Product: glycoprotein B #status predicted <GPB>
F:740-809/Domain: transmembrane #status predicted <TMN>
F:106,1216,321,364,438,456,493,499,666,688/Binding site: carbohydrate (Asn) (covalent)
Query Match 26.4%; Score 86; DB 1; Length 919;
Best Local Similarity 35.0%; Pred. No. 0.061;
Matches 21; Conservative 7; Mismatches 14; Indels 18; Gaps 2;
QY 20 NSSPST-----QNTVSREVSSVQLSEES-----TFYLCPPVPGSVIRLE 61
DB 34 SSQSTPASTQSAKTVDTLLTETPDPLRLAVRESGILAEEDGDFYTCPPPTGATVVRLE 93
RESULT 15
T42576
probable envelope protein 33 - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42576
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.

A:Reference number: 222173; MUID:98264497; PMID:9603335
A:Accession: T42576
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-975 <TEL>
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59549.1; PID:g2605977
A:Experimental source: strain NS80567
C:Genetics:
A:Gene: 33
C:Superfamily: herpesvirus glycoprotein B

Query Match 26.4%; Score 86; DB 2; Length 975;
Best Local Similarity 35.0%; Pred. No. 0.065; 14; Indels 18; Gaps 2;
Matches 21; Conservative 7; Mismatches 14; Indels 18; Gaps 2;

QY 20 NSSPST-----QNVTSREVVSVQLSEES-----TFYLCPPPVGVSTVIRLE 61
: ||| : : : : : ||| ||| : : : : :
Db 90 SSQPSTPASTQSAKTVDQTLTETPDPLRLAVRESGILAEDEGDFYTCPPPTGSTVVRIE 149

Search completed: October 8, 2003, 17:08:24
Job time : 3.1581 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	85	26.1	904	1	PCT-US03-11231-18		Sequence 18, Appl
2	61	18.7	1399	1	PCT-US02-21485-26		Sequence 26, Appl
3	60	18.4	1671	1	PCT-US02-21485-18		Sequence 18, Appl
4	58.5	17.9	294	6	US-10-425-114A-39748		Sequence 39748, A
5	58.5	17.9	294	6	US-10-425-114A-39755		Sequence 39755, A
6	58	17.8	1471	1	PCT-US02-21485-20		Sequence 20, Appl
7	57	17.5	257	5	US-10-425-114A-53240		Sequence 53240, A
8	57	17.5	258	5	US-09-976-858-5		Sequence 5, Appl
9	57	17.5	258	5	US-09-976-858-6		Sequence 6, Appl
10	57	17.5	1438	1	PCT-US02-21485-22		Sequence 22, Appl
11	57	17.5	1578	1	PCT-US02-21485-24		Sequence 24, Appl
12	56.5	17.3	181	1	PCT-US03-26780-1638		Sequence 1638, Ap
13	56.5	17.3	479	1	PCT-US03-28227-4020		Sequence 4020, Ap
14	56.5	17.3	504	7	US-60-499-838-8		Sequence 8, Appl
15	56	17.2	239	1	PCT-US02-06765-85		Sequence 85, Appl
16	55.5	17.0	137	1	PCT-US03-20322-225		Sequence 225, App
17	55	16.9	207	6	US-10-425-114A-51239		Sequence 51239, A
18	54.5	16.7	432	1	PCT-US02-36090-19		Sequence 19, Appl
19	54.5	16.7	767	1	PCT-US02-36090-44		Sequence 44, Appl
20	54.5	16.7	1063	1	PCT-US02-36090-43		Sequence 43, Appl
21	54.5	16.7	1241	1	PCT-US02-36090-17		Sequence 17, Appl
22	54	16.6	1885	1	PCT-US03-28626-23		Sequence 23, Appl
23	54	16.6	1885	6	US-10-661-398-23		Sequence 23, Appl
24	53.5	16.4	884	1	PCT-US03-28227-5394		Sequence 5394, Ap
25	53.5	16.4	889	1	PCT-US03-28227-5393		Sequence 5393, Ap
26	53.5	16.4	891	1	PCT-US03-28227-5392		Sequence 5392, Ap

```

Db      1330 NRSTASIQPTSDLLVSAECSSDDDEDFVECEPSTGGELV 1368
          | | :: ||::||| : | :: | | | | | | ::
RESULT 4
US-10-425-114A-39748
; Sequence 39748, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39748
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700725110_FLI.pep
US-10-425-114A-39748

Query Match              17.9%; Score 58.5; DB 6; Length 294;
Best Local Similarity    25.9%; Pred. No. 32;
Matches 14; Conservative 11; Mismatches 20; Indels 9; Gaps 1;

QY      3 YFRNCIFFLIVLYGTNSSPSTONVTISREVSVVOLSEESTFVLCPPPVGST 56
        ||:::||::||| : | :: | | | | | | ::
DB      166 FFNQTDVHSIDVFPFTALSTGNTSTTTTTSQI-----LVCPPPPGAT 210

RESULT 5
US-10-425-114A-39755
; Sequence 39755, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39755
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700729312_FLI.pep
US-10-425-114A-39755

Query Match              17.9%; Score 58.5; DB 6; Length 294;
Best Local Similarity    25.9%; Pred. No. 32;
Matches 14; Conservative 11; Mismatches 20; Indels 9; Gaps 1;

QY      3 YFRNCIFFLIVLYGTNSSPSTONVTISREVSVVOLSEESTFVLCPPPVGST 56
        ||:::||::||| : | :: | | | | | | ::
DB      166 FFNQTDVHSIDVFPFTALSTGNTSTTTTTSQI-----LVCPPPPGAT 210

RESULT 6
PCT-US02-21485-20
; Sequence 20, Application PC/TUS0221485

```



```

QY      5 RNCRTFF----LIVILYGNTSSPSTONVTSREVWSSVOLSEESTFYLCPPPVGVSTVIRL 60
       | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     165 RLHCFFYFPSSIPPCLYASSSTPEQSNVHDARHVFDRGRIDSSSFFIIRDITASFCCTMRL 224

RESULT 8
US-09-976-858-5
; Sequence 5, Application US/09976858
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Mack, David H.
; APPLICANT: Wilson, Keith E.
; APPLICANT: Afar, Daniel
; APPLICANT: Peter, Hevezi
; TITLE OF INVENTION: Methods of Diagnosis of Prostate Cancer, Compositions and Meth
; FILE REFERENCE: 05882.0183.NPUS00
; CURRENT APPLICATION NUMBER: US/09/976,858
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/276,791
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/288,589
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/276,888
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/286,214
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/281,922
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/263,957
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 258
; TYPE: PRT
; ORGANISM: human organism
US-09-976-858-5

Query Match          17.5%; Score 57; DB 5; Length 258;
Best Local Similarity 36.1%; Pred.No.40;
Matches 13; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY      8 CIFFLVILYGNTSSPSTONVTSREVWSSVOLSEEE 43
       | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     17 CEASTVLLNSTDSSPTNFTDIEAALKAQLDSAD 52

RESULT 9
US-09-976-858-6
; Sequence 6, Application US/09976858
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Mack, David H.
; APPLICANT: Wilson, Keith E.
; APPLICANT: Afar, Daniel
; APPLICANT: Peter, Hevezi
; TITLE OF INVENTION: Methods of Diagnosis of Prostate Cancer, Compositions and Meth
; FILE REFERENCE: 05882.0183.NPUS00
; CURRENT APPLICATION NUMBER: US/09/976,858
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/276,791
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/288,589
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/276,888
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/286,214
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/281,922
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/263,957

```



```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1638
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-26780-1638

Query Match      17.3%; Score 56.5; DB 1; Length 181;
Best Local Similarity 29.6%; Pred. No. 34;
Matches 16; Conservative 10; Mismatches 25; Indels 3; Gaps 1;

QY 11 FLVILYGTNSPSTQNTSREVSVVLS---EEESTFYLCPPVPGSTVIRLE 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 FEAVLRVEEERANSQNTIKREVEDDGLSLMIDSONNQYILTKPRDSTIPRAD 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
PCT-US03-28227-4020
; Sequence 4020, Application PC/TUS0328227
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANYILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: PANZER, Kristian A.; BLANCHARD, John L.;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 4020
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 950149.PT14p
PCT-US03-28227-4020

Query Match      17.3%; Score 56.5; DB 1; Length 479;
Best Local Similarity 30.5%; Pred. No. 75;
Matches 18; Conservative 8; Mismatches 22; Indels 11; Gaps 3;

QY 4 FRRNC-IFPLIVILY-----GTNSSPSTQNTSREVW-SSVOLSEESTFYLCPP 51
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 FSLSCWVFELIAVIYKHFVPCPLPPNNTTGNFSHVEIVKEVQLQVEPEASAFCTP 252
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-60-499-838-8
; Sequence 8, Application US/60499838
```

```
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; FILE REFERENCE: CARDIOVASCULAR DISEASE USING 554, 16408, 42028 OR 11209
; CURRENT APPLICATION NUMBER: US/60/499,838
; CURRENT FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-60-499-838-8

Query Match      17.3%; Score 56.5; DB 7; Length 504;
Best Local Similarity 30.5%; Pred. No. 78;
Matches 18; Conservative 8; Mismatches 22; Indels 11; Gaps 3;

QY 4 FRRNC-IFPLIVILY-----GTNSSPSTQNTSREVW-SSVOLSEESTFYLCPP 51
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 FSLSCWVFELIAVIYKHFVPCPLPPNNTTGNFSHVEIVKEVQLQVEPEASAFCTP 277
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
PCT-US02-06765-85
; Sequence 85, Application PC/TUS0206765
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: ALK Abello
; TITLE OF INVENTION: RECOMBINANT HYBRID ALLERGEN CONSTRUCTS WITH REDUCED ALLERGENIC
; FILE REFERENCE: 2313/2H587WO0
; CURRENT APPLICATION NUMBER: PCT/US02/06765
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/272,818
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-06765-85

Query Match      17.2%; Score 56; DB 1; Length 239;
Best Local Similarity 38.7%; Pred. No. 48;
Matches 12; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 13 IVILYGTNSPSTQNTSREVSVVOLSEEE 43
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 VVLLNSTDSSPPTNNFTDIEAALKALQDSAD 33
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: October 8, 2003, 17:25:44
Job time : 1.18972 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:00 ; Search time 25.2332 Seconds
(without alignments)
2271.804 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFLVLVLYGTN.....STFYLCPPPVGVSTVIRLEFG 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA.Main:*

- 1: /cgn2_6/ptodata/1/paa/PTCUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
- 32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326	100.0	456	15	US-09-147-052-2 Sequence 2, Appl1

2	326	100.0	456	24	US-09-901-572A-3	Sequence 3, Appl1
3	320	98.2	62	27	US-10-131-591A-5	Sequence 5, Appl1
4	314	96.3	865	3	US-07-722-860-13	Sequence 13, Appl1
5	314	96.3	865	6	US-08-213-449A-13	Sequence 13, Appl1
6	314	96.3	865	6	US-08-213-449B-13	Sequence 13, Appl1
7	314	96.3	1086	15	US-09-147-052-4	Sequence 4, Appl1
8	308	94.5	62	27	US-10-131-591A-6	Sequence 6, Appl1
9	107	32.8	891	9	US-08-541-878-6	Sequence 8, Appl1
10	100	30.7	868	12	US-08-804-439-21	Sequence 23, Appl1
11	100	30.7	868	17	US-09-301-390-21	Sequence 21, Appl1
12	100	30.7	868	17	US-09-338-326-21	Sequence 21, Appl1
13	99.5	30.5	250	19	US-09-521-738-16	Sequence 16, Appl1
14	99.5	30.5	943	4	US-08-096-183D-4	Sequence 4, Appl1
15	99.5	30.5	943	9	US-09-521-738-14	Sequence 14, Appl1
16	96	29.4	885	9	US-08-541-878-8	Sequence 8, Appl1
17	96	29.4	885	12	US-08-804-439-23	Sequence 23, Appl1
18	96	29.4	885	17	US-09-301-390-23	Sequence 23, Appl1
19	96	29.4	885	17	US-09-338-326-23	Sequence 23, Appl1
20	95	29.1	903	12	US-08-804-439-22	Sequence 22, Appl1
21	95	29.1	903	17	US-09-301-390-22	Sequence 22, Appl1
22	95	29.1	903	17	US-09-338-326-22	Sequence 22, Appl1
23	95	29.1	904	32	US-60-412-956-12	Sequence 12, Appl1
24	87	26.7	200	22	US-09-791-537-145956	Sequence 145956,
25	85	26.1	846	5	US-08-123-456-142	Sequence 142, App
26	85	26.1	846	16	US-09-297-477A-142	Sequence 142, App
27	85	26.1	846	25	US-09-994-404-142	Sequence 142, App
28	85	26.1	896	5	US-08-123-456-241	Sequence 241, App
29	85	26.1	896	16	US-09-297-477A-241	Sequence 241, App
30	85	26.1	896	25	US-09-994-404-241	Sequence 241, App
31	85	26.1	904	1	PCR-US03-11231-18	Sequence 18, Appl1
32	85	26.1	904	27	US-10-121-988-18	Sequence 18, Appl1
33	85	26.1	904	28	US-10-200-562-18	Sequence 18, Appl1
34	85	26.1	904	28	US-10-237-551-18	Sequence 18, Appl1
35	82.5	25.3	933	3	US-07-805-524-2	Sequence 2, Appl1
36	64	19.6	109	30	US-10-437-963-169354	Sequence 169354,
37	64	19.6	1604	32	US-60-173-464-28691	Sequence 28691, A
38	64	19.6	1605	32	US-09-614-150-37917	Sequence 37917, A
39	64	19.6	1605	32	US-60-191-637-37534	Sequence 37534, A
40	64	19.6	1605	32	US-60-191-681-29173	Sequence 29173, A
41	63	19.3	639	22	US-09-791-537-77065	Sequence 77065, A
42	63	19.3	649	22	US-09-791-537-30976	Sequence 30976, A
43	63	19.3	649	22	US-09-791-537-144118	Sequence 144118,
44	63	19.3	1395	22	US-09-791-537-114414	Sequence 114414,
45	63	19.3	1395	22	US-09-791-537-144605	Sequence 144605,

ALIGNMENTS

RESULT 1

US-09-147-052-2

; Sequence 2, Application US/09147052

; GENERAL INFORMATION:

; APPLICANT: SAITOH, Shuji

; APPLICANT: TSUZAKI, Yoshinari

; APPLICANT: YANAGIDA, Noboru

; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,

; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE

; FILE REFERENCE: 981167

; CURRENT APPLICATION NUMBER: US/09/147,052

; CURRENT FILING DATE: 1999-04-05

; PRIOR FILING DATE: 1996-03-29

; PRIOR APPLICATION NUMBER: JP 08-103548

; PRIOR APPLICATION NUMBER: PCT/JP97/01084

; PRIOR FILING DATE: 1997-03-28

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 456

; TYPE: PRT

; ORGANISM: hybrid

US-09-147-052-2

```
Query Match      100.0%; Score 326; DB 15; Length 456;
Best Local Similarity 100.0%; Pred. No. 5.7e-32;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 EFG 63
DB 61 EFG 63

RESULT 2
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901.572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pn240K-S
US-09-901-572A-3

Query Match      100.0%; Score 326; DB 24; Length 456;
Best Local Similarity 100.0%; Pred. No. 5.7e-32;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 EFG 63
DB 61 EFG 63

RESULT 3
US-10-131-591A-5
; Sequence 5, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: MDVgB signal
US-10-131-591A-5

Query Match      98.2%; Score 320; DB 27; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.9e-32;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 EF 62
```

```
DB 61 EF 62

RESULT 4
US-07-722-860-13
; Sequence 13, Application US/07722860
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, YI
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 North Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/722.860
; FILING DATE: 19910628
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-101P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-722-860-13

Query Match      96.3%; Score 314; DB 3; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.2e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTSREVSVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 E 61
DB 61 E 61

RESULT 5
US-08-213-449A-13
; Sequence 13, Application US/08213449A
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, YI
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
```



```

; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified VgB signal
US-10-131-591A-6

Query Match          94.5%; Score 308; DB 27; Length 62;
Best Local Similarity 96.8%; Pred. No. 9,7e-31;
Matches 60; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYKNTSPQNTSVSSVOLSEESTFYLCPPPVGTVIRL 60
    |||||
Db 1 MHYFRNCIFFLIVLYKNTSPQNTSVSSVOLSEESTFYLCPPPVGTVIRL 60
    |||||
QY 61 EF 62
    ||
Db 61 EF 62

RESULT 9
US-08-541-878-6
; Sequence 6, Application US/08541878
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Black, Darla
; APPLICANT: Scinicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cox & Smith Incorporated
; STREET: 112 East Pecan Street, Suite 2000
; CITY: San Antonio
; STATE: Texas
; COUNTRY: USA
; ZIP: 78205
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,878
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/042,747
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haymond, W. Bradley
; REGISTRATION NUMBER: 35186
; REFERENCE/DOCKET NUMBER: S-0072.179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 210-554-5500
; TELEFAX: 210-226-8395
; TELEX: 767609
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 891 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-541-878-6

Query Match          32.8%; Score 107; DB 9; Length 891;
Best Local Similarity 40.9%; Pred. No. 0.00088;
Matches 18; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

```

```

QY 18 GTNSPSTQNTSVSSVOLSEESTFYLCPPPVGTVIRLE 61
    |||
Db 72 GTNASVEAGHATLRENLRDIALDGDATFYVCPPTGATVYQFE 115
    |||

RESULT 10
US-08-804-439-21
; Sequence 21, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439-21

Query Match          30.7%; Score 100; DB 12; Length 868;
Best Local Similarity 54.8%; Pred. No. 0.0066;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 31 REVSVSSVOLSEESTFYLCPPPVGTVIRLE 61
    |||
Db 41 REAIHKSDAETKPTFYVCPPTGTVIRLE 71
    |||

RESULT 11
US-09-301-390-21
; Sequence 21, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

```



```
;
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-301-390-21

Query Match          30.7%; Score 100; DB 17; Length 868;
Best Local Similarity 54.8%; Pred. No. 0.0066;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 31 REVSSVQLSEESTFYLCPPPVGVSTVIRLE 61
Db 41 REAIHKSQDAETKPTFFVCPPTGVSTVIRLE 71

RESULT 12
US-09-338-326-21
; Sequence 21, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-338-326-21

Query Match          30.7%; Score 100; DB 17; Length 868;
Best Local Similarity 54.8%; Pred. No. 0.0066;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 31 REVSSVQLSEESTFYLCPPPVGVSTVIRLE 61
Db 41 REAIHKSQDAETKPTFFVCPPTGVSTVIRLE 71

RESULT 13
US-09-521-738-16
; Sequence 16, Application US/09521738
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9
; CURRENT APPLICATION NUMBER: US/09/521.738
; CURRENT FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
; US-09-521-738-16

Query Match          30.5%; Score 99.5; DB 19; Length 250;
Best Local Similarity 24.6%; Pred. No. 0.0017;
Matches 29; Conservative 14; Mismatches 16; Indels 59; Gaps 4;

Qy 3 YFRRCIF-----FLIVLY-----GTNSSP 23
Db 22 YFRCFFPSLLGIATNGSRHNGSSGLTRLARVSVFIWVFLVGRPRVPGSGSTSEQ 81
Qy 24 STQNVTSREV-----VSSVOLSEEE-----STFYLCPPPVGVSTVIRLE 61
Db 82 PRRTVATPEGVGHQNLQIPPICRYEALRASQIEANGPSTFFYMCPPPGSGSTVIRLE 139

RESULT 14
US-08-096-183D-4
; Sequence 4, Application US/08096183D
; GENERAL INFORMATION:
; APPLICANT: Roger K. Maes and Stephen J. Spatz
; TITLE OF INVENTION: Recombinant Poxvirus
; TITLE OF INVENTION: Vaccine Against
; TITLE OF INVENTION: Feline Rhinotracheitis
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 KB
```

```
/ MEDIUM TYPE: storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/096,183D
/ FILING DATE: July 26, 1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ian C. McLeod
/ REGISTRATION NUMBER: 20,931
/ REFERENCE/DOCKET NUMBER: MSU 4.1-166
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (517) 347-4100
/ TELEFAX: (517) 347-4103
/ TELEX: None
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 943
/ TYPE: Amino Acid
/ STRANDEDNESS: Single
/ TOPOLOGY: Linear
/ MOLECULE TYPE:
/ DESCRIPTION: Polypeptide
/ HYPOTHETICAL: No
/ ANTI-SENSE: No
/ ORIGINAL SOURCE:
/ ORGANISM: Feline herpesvirus-1
/ STRAIN: 1
/ INDIVIDUAL ISOLATE: C-27
/ CELL TYPE: N/A
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD: Deduced sequence
/ OTHER INFORMATION: gb
US-08-096-183D-4

Query Match 30.5%; Score 99.5; DB 4; Length 943;
Best Local Similarity 24.6%; Pred. No. 0.0085;
Matches 29; Conservative 14; Mismatches 16; Indels 59; Gaps 4;

QY 3 YFRRNCIP-----FLIVLY-----GTNSSP 23
DB 22 YFQRCFFPSLLGIAATGSRHNGSSGLRLARYVFIWVLFVLPVPVEGSGSTSEQ 81
QY 24 STQNVTSREV-----VSSVOLSEEE---STFYLCPPPPVGGSTVIRLE 61
DB 82 PRRTVATPEGVGHQNLQIPPICRYEALRASQIEANGPSTFYMCPPPSGSTVVRLE 139

Query Match 30.5%; Score 99.5; DB 19; Length 943;
Best Local Similarity 24.6%; Pred. No. 0.0085;
Matches 29; Conservative 14; Mismatches 16; Indels 59; Gaps 4;

QY 3 YFRRNCIP-----FLIVLY-----GTNSSP 23
DB 22 YFQRCFFPSLLGIAATGSRHNGSSGLRLARYVFIWVLFVLPVPVEGSGSTSEQ 81
QY 24 STQNVTSREV-----VSSVOLSEEE---STFYLCPPPPVGGSTVIRLE 61
DB 82 PRRTVATPEGVGHQNLQIPPICRYEALRASQIEANGPSTFYMCPPPSGSTVVRLE 139

Search completed: October 8, 2003, 17:24:54
Job time : 26.2332 secs
```

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM protein - protein search, using sw model
 Run on: October 8, 2003, 17:07:16 ; Search time 3.62451 Seconds
 (without alignments)
 2800.682 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63
 Perfect score: 326
 Sequence: 1 MHVFRNCIFLLVILYGTN.....STFYLCPPPGVSTVIRLEFG 63

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications_AA:
 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	326	100.0	456	9	US-09-147-052-2
2	326	100.0	456	12	US-09-901-572A-3
3	320	98.2	62	15	US-10-131-591A-5
4	314	96.3	1086	9	US-09-147-052-4
5	308	94.5	62	15	US-10-131-591A-6
6	85	26.1	904	9	US-09-894-998-18
7	85	26.1	904	12	US-10-200-562-18
8	85	26.1	904	12	US-10-237-551-18
9	85	26.1	904	15	US-10-121-988-18
10	62	19.0	1284	15	US-10-251-385-294
11	60.5	18.6	444	8	US-08-812-393A-2
12	59	18.1	1435	16	US-10-282-287-8
13	59	18.1	1436	9	US-09-815-242-5566
14	59	18.1	1442	9	US-09-815-242-12321
15	57	17.5	257	11	US-09-800-198-88
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 4, Appli
					Sequence 6, Appli
					Sequence 18, Appl
					Sequence 18, Appl
					Sequence 18, Appl
					Sequence 294, Appl
					Sequence 2, Appli
					Sequence 8, Appli
					Sequence 5566, Ap
					Sequence 12321, A
					Sequence 88, Appl

16	57	17.5	258	10	US-09-808-602-110	Sequence 110, App
17	57	17.5	258	11	US-09-800-198-96	Sequence 96, Appl
18	56.5	17.3	481	10	US-09-815-923-8	Sequence 8, Appli
19	56.5	17.3	504	10	US-09-919-497-67	Sequence 67, Appli
20	56.5	17.3	504	12	US-10-044-897-9	Sequence 9, Appli
21	56.5	17.3	504	12	US-10-044-901-7	Sequence 7, Appli
22	56.5	17.3	524	11	US-09-746-783-13	Sequence 13, Appli
23	56.5	17.3	677	12	US-10-137-870-230	Sequence 230, App
24	56.5	17.3	677	12	US-10-140-018-230	Sequence 230, App
25	56.5	17.3	677	12	US-10-140-021-230	Sequence 230, App
26	56.5	17.3	677	12	US-10-140-274-230	Sequence 230, App
27	56.5	17.3	677	12	US-10-140-471-230	Sequence 230, App
28	56.5	17.3	677	12	US-10-140-807-230	Sequence 230, App
29	56.5	17.3	677	12	US-10-140-922-230	Sequence 230, App
30	56.5	17.3	677	12	US-10-140-924-230	Sequence 230, App
31	56.5	17.3	677	12	US-10-140-926-230	Sequence 230, App
32	56.5	17.3	677	12	US-10-141-698-230	Sequence 230, App
33	56.5	17.3	677	12	US-10-141-702-230	Sequence 230, App
34	56.5	17.3	677	12	US-10-141-704-230	Sequence 230, App
35	56.5	17.3	677	12	US-10-142-421-230	Sequence 230, App
36	56.5	17.3	677	12	US-10-142-432-230	Sequence 230, App
37	56.5	17.3	677	12	US-10-142-767-230	Sequence 230, App
38	56.5	17.3	677	12	US-10-143-033-230	Sequence 230, App
39	56.5	17.3	677	12	US-10-144-994-230	Sequence 230, App
40	56.5	17.3	677	12	US-10-145-628-230	Sequence 230, App
41	56.5	17.3	677	12	US-10-145-633-230	Sequence 230, App
42	56.5	17.3	677	12	US-10-145-633-230	Sequence 230, App
43	56.5	17.3	677	12	US-10-145-746-230	Sequence 230, App
44	56.5	17.3	677	12	US-10-145-748-230	Sequence 230, App
45	56.5	17.3	677	12	US-10-145-823-230	Sequence 230, App

ALIGNMENTS

RESULT 1
 US-09-147-052-2
 ; Sequence 2, Application US/09147052
 ; Patent No. US20010014335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SAITOH, Shuji
 ; APPLICANT: TSUZAKI, Yoshinari
 ; APPLICANT: YANAGIDA, NO. US20010014335A1
 ; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
 ; FILE REFERENCE: 981167
 ; CURRENT APPLICATION NUMBER: US/09/147,052
 ; PRIOR FILING DATE: 1999-04-05
 ; PRIOR APPLICATION NUMBER: JP 08-103548
 ; PRIOR FILING DATE: 1998-03-29
 ; PRIOR APPLICATION NUMBER: PCT/JP97/01084
 ; PRIOR FILING DATE: 1997-03-28
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: hybrid
 US-09-147-052-2

Query Match 100.0%; Score 326; DB 9; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-32;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHVFRNCIFLLVILYGTN...STFYLCPPPGVSTVIRLEFG 60
 |||
 DB 1 MHVFRNCIFLLVILYGTN...STFYLCPPPGVSTVIRLEFG 60
 |||

QY 61 EFG 63
 |||
 DB 61 EFG 63

```
RESULT 2
US-09-901-572A-3
; Sequence 3, Application US/09901572A
; Publication No. US2003016534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901.572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3
Query Match          100.0%; Score 326; DB 12; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.5e-32;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 EFG 63
    |||
Db 61 EFG 63

RESULT 3
US-10-131-591A-5
; Sequence 5, Application US/10131591A
; Publication No. US2003005979A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: MDVgB signal
US-10-131-591A-5
Query Match          98.2%; Score 320; DB 15; Length 62;
Best Local Similarity 100.0%; Pred. No. 6.9e-33;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 EF 62
    ||
Db 61 EF 62

RESULT 4
US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: MCGOWAN, Patrick

; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147.052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4
Query Match          96.3%; Score 314; DB 9; Length 1086;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 E 61
    |
Db 61 E 61

RESULT 5
US-10-131-591A-6
; Sequence 6, Application US/10131591A
; Publication No. US2003005979A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: Modified VgB signal
US-10-131-591A-6
Query Match          94.5%; Score 308; DB 15; Length 62;
Best Local Similarity 96.8%; Pred. No. 2.2e-31;
Matches 60; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MHYFRNCIFFLIVILYGTNSPSTQNTVSREVVSVOLSEESTFYLCPPVPGSTVIRL 60
QY 61 EF 62
    ||
Db 61 EF 62

RESULT 6
US-09-894-998-18
; Sequence 18, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: MCGOWAN, Patrick
```

US-10-237-551-18

QY 2 HYFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVQLSEESTFYL 48

Db 1117 HNELVQCSFFLC-----GVNSKASKSSITKHSMTOKTQFRKVLGVFYL 11159

RESULT 11

US-08-812-393A-2

; Sequence 2, Application US/08812393A

; Publication No. US20010007152A1

; GENERAL INFORMATION:

; APPLICANT: SHERMAN, Linda A.

; APPLICANT: LUSTGARTEN, Joseph

; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING

; TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR

; TITLE OF INVENTION: ANTIGENS

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Avenue, NW, suite 5500

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006-1888

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA: US/08/812.393A

; FILING DATE: 05-MAR-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Murashige, Kate H

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 31333-20001.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-887-1500

; TELEFAX: 202-822-0168

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 444 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-08-812-393A-2

Query Match 18.6%; Score 60.5; DB 8; Length 444;

Best Local Similarity 28.8%; Pred. No. 34;

Matches 17; Conservative 12; Mismatches 23; Indels 7; Gaps 2;

QY 12 LIVLGYTNSPSTQNTVTSREVY----SSVOLSEEST---FYLCPVPVGVSTVIRLEFG 63

Db 65 LVSELLNPGSKSGRLTSTTVIKERRSSLUHSSQITDSGTLYLCASNSGGSNAKLTFG 123

RESULT 12

US-10-282-287-8

; Sequence 8, Application US/10282287

; Publication No. US20030129633A1

; GENERAL INFORMATION:

; APPLICANT: O'Donnell, Michael E.

; APPLICANT: Zhang, Dan

; APPLICANT: Whipple, Richard

; TITLE OF INVENTION: DNA REPLICATION PROTEINS OF GRAM POSITIVE BACTERIA AND

; TITLE OF INVENTION: THEIR USE TO SCREEN FOR CHEMICAL INHIBITORS

; FILE REFERENCE: 22221/1002

; CURRENT APPLICATION NUMBER: US/10/282,287

; CURRENT FILING DATE: 2002-10-28

; PRIOR APPLICATION NUMBER: US/09/235,245

; PRIOR FILING DATE: 1999-01-22

; PRIOR APPLICATION NUMBER: 60/074,522

; PRIOR FILING DATE: 1998-01-27

; PRIOR APPLICATION NUMBER: 60/093,727

; PRIOR FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 1435

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-10-282-287-8

Query Match 18.1%; Score 59; DB 9; Length 1436;

Best Local Similarity 35.9%; Pred. No. 2.3e+02;

Matches 14; Conservative 9; Mismatches 14; Indels 2; Gaps 1;

QY 6 RNCIFFLIVLYGTNSPSTQNTVTSREVSVSVLSSEES 44

Db 149 RNCGFIDKIFETNDNQNLASLE--AHIQEEDQS 185

RESULT 13

US-09-815-242-5566

; Sequence 5566, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5566

; LENGTH: 1436

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5566

Query Match 18.1%; Score 59; DB 9; Length 1436;

Best Local Similarity 35.9%; Pred. No. 2.3e+02;

Matches 14; Conservative 9; Mismatches 14; Indels 2; Gaps 1;

QY 6 RNCIFFLIVLYGTNSPSTQNTVTSREVSVSVLSSEES 44

Db 149 RNCGFIDKIFETNDNQNLASLE--AHIQEEDQS 185

RESULT 14

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:40 ; Search time 1.77075 Seconds
(without alignments)
1505.341 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFFLVILYGIN.....STFYLCPPPVGVSTVIRLFG 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	314	96.3	865	1	US-07-803-633A-13
2	107.5	33.0	879	1	US-08-220-151-2
3	107.5	33.0	879	1	US-08-220-151-3
4	107.5	33.0	879	1	US-08-413-118-2
5	107.5	33.0	879	1	US-08-413-118-3
6	107.5	33.0	879	1	US-08-413-118-106
7	107.5	33.0	879	3	US-08-473-446-2
8	107.5	33.0	879	3	US-08-473-446-3
9	107.5	33.0	879	3	US-08-473-446-106
10	107	32.8	891	1	US-08-042-747A-6
11	106	32.5	1041	1	US-08-220-151-4
12	106	32.5	1041	1	US-08-413-118-4
13	106	32.5	1041	3	US-08-473-446-4
14	100	30.7	868	1	US-08-220-151-7
15	100	30.7	868	1	US-08-413-118-7
16	100	30.7	868	3	US-08-804-439A-21
17	100	30.7	868	3	US-08-473-446-7
18	100	30.7	868	3	US-08-720-229-21
19	99.5	30.5	943	3	US-08-911-321-4
20	96	29.4	885	1	US-08-042-747A-8
21	96	29.4	885	3	US-08-804-439A-23
22	96	29.4	885	3	US-08-720-229-23
23	95	29.1	903	1	US-08-220-151-8
24	95	29.1	903	1	US-08-413-118-8
25	95	29.1	903	3	US-08-804-439A-22
26	95	29.1	903	3	US-08-473-446-8
27	95	29.1	903	3	US-08-720-229-22

28 95 29.1 904 3 US-08-632-537-1 Sequence 1, Appli
29 95 29.1 904 5 PCT-US96-05316-1 Sequence 1, Appli
30 95 29.1 904 6 5244792-4 Patent No. 5244792
31 91 27.9 980 1 US-08-220-151-5 Sequence 5, Appli
32 91 27.9 980 1 US-08-413-118-5 Sequence 5, Appli
33 91 27.9 980 3 US-08-473-446-5 Sequence 5, Appli
34 85 26.1 904 3 US-08-632-537-2 Sequence 2, Appli
35 85 26.1 904 4 US-09-894-998A-18 Sequence 18, Appli
36 85 26.1 904 5 PCT-US96-05316-2 Sequence 2, Appli
37 85 26.1 904 6 5244792-3 Patent No. 5244792
38 84 25.8 913 1 US-08-220-151-6 Sequence 6, Appli
39 84 25.8 913 1 US-08-413-118-6 Sequence 6, Appli
40 84 25.8 913 3 US-08-473-446-6 Sequence 6, Appli
41 84 25.8 913 3 US-09-232-468A-2 Sequence 2, Appli
42 84 25.8 913 4 US-09-784-984B-49 Sequence 49, Appli
43 84 25.8 913 6 5196516-8 Patent No. 5196516
44 82.5 25.3 933 2 US-08-682-847-2 Sequence 2, Appli
45 62 19.0 1284 4 US-09-170-496D-294 Sequence 294, App

ALIGNMENTS

RESULT 1
US-07-803-633A-13
; Sequence 13, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/803.633A
; APPLICATION NUMBER: 19911210
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-803-633A-13

Query Match 96.3%; Score 314; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILYGTNTPSTQNTVTSREVSSVOLSEESTFYLCPPPVGVSTVIRL 60

```

Db      1 MHYFRNCIFFLIVLYGNTSSPSTQNTSREVSVSQLSEESTFYLCPPVGSTVIRL 60
QY      61 E 61
Db      61 E 61

RESULT 2
US-08-220-151-2
; Sequence 2, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-3333
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-220-151-2

Query Match      33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY      9 IFFLVIL-----YGTNSSPSTQNTSREVSVSQLSEESTFYLCPP 51
Db      8 IFFIYLIICDPTTPESTINPLNHNLSLTPKPTSDDIKREILRESQIESDDTSTFYMCPP 67

QY      52 PVGSTVIRLE 61
Db      68 PSGSTLVRLE 77

RESULT 3
US-08-220-151-3
; Sequence 3, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118

```

```

; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-3

Query Match      33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY      9 IFFLVIL-----YGTNSSPSTQNTSREVSVSQLSEESTFYLCPP 51
Db      8 IFFIYLIICDPTTPESTINPLNHNLSLTPKPTSDDIKREILRESQIESDDTSTFYMCPP 67

QY      52 PVGSTVIRLE 61
Db      68 PSGSTLVRLE 77

RESULT 4
US-08-413-118-2
; Sequence 2, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118

```

```
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-413-118-2

Query Match          33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY 9 IFFLLIVL-----YGTNSSPSTQNTVTSREVSVVOL-SEESTFYLCPP 51
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 8 IFFIITLIICDPTTPESTINPLNHHNLSTPKPTSDIREILRESQIESDDTSTFYMCPP 67
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
QY 52 PVGSTVIRLE 61
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 68 PSGSTLVRL 77
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 5
US-08-413-118-3
; Sequence 3, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-413-118-3

Query Match          33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY 9 IFFLLIVL-----YGTNSSPSTQNTVTSREVSVVOL-SEESTFYLCPP 51
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 8 IFFIITLIICDPTTPESTINPLNHHNLSTPKPTSDIREILRESQIESDDTSTFYMCPP 67
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
QY 52 PVGSTVIRLE 61
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 68 PSGSTLVRL 77
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
```

```
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-413-118-3

Query Match          33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY 9 IFFLLIVL-----YGTNSSPSTQNTVTSREVSVVOL-SEESTFYLCPP 51
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 8 IFFIITLIICDPTTPESTINPLNHHNLSTPKPTSDIREILRESQIESDDTSTFYMCPP 67
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
QY 52 PVGSTVIRLE 61
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 68 PSGSTLVRL 77
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 6
US-08-413-118-106
; Sequence 106, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-413-118-106

Query Match          33.0%; Score 107.5; DB 1; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY 9 IFFLLIVL-----YGTNSSPSTQNTVTSREVSVVOL-SEESTFYLCPP 51
   ||||| : | : | : | : | : | : | : | : | : | : | : | : |
```

```
Db      8 IFFIYTLICDPTTPESTINPLNHNHLSKPSTDDIREILRESQIESDDTSTFYMCPP 67
QY      52 PVGSTVIRLE 61
        | : : : : |
Db      68 PSGSTLVRL 77
        | : : : : |

RESULT 7
US-08-473-446-2
; Sequence 2, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
US-08-473-446-2

Query Match          33.0%; Score 107.5; DB 3; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY      9 IFFLVIL-----YGTNSPSTQNTSREVSVVOL-SEESTFYLCPP 51
        | : : : |
Db      8 IFFIYTLICDPTTPESTINPLNHNHLSKPSTDDIREILRESQIESDDTSTFYMCPP 67
        | : : : |

QY      52 PVGSTVIRLE 61
        | : : : : |
Db      68 PSGSTLVRL 77
        | : : : : |

RESULT 8
US-08-473-446-3
; Sequence 3, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-473-446-3

Query Match          33.0%; Score 107.5; DB 3; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;

QY      9 IFFLVIL-----YGTNSPSTQNTSREVSVVOL-SEESTFYLCPP 51
        | : : : |
Db      8 IFFIYTLICDPTTPESTINPLNHNHLSKPSTDDIREILRESQIESDDTSTFYMCPP 67
        | : : : |

QY      52 PVGSTVIRLE 61
        | : : : : |
Db      68 PSGSTLVRL 77
        | : : : : |

RESULT 9
US-08-473-446-106
; Sequence 106, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/473.446
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/413.118
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 879 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-473-446-106

Query Match 33.0%; Score 107.5; DB 3; Length 879;
Best Local Similarity 35.7%; Pred. No. 1.2e-05;
Matches 25; Conservative 12; Mismatches 16; Indels 17; Gaps 2;
QY 9 IFFLVIL-----YGTNSPSTQNTVTSREVSVVOL-SEESTFYLCPP 51
DB 8 IFFIYLIICDPTTPESTINPLNHNLSLTPKPSDDIREILRESQIESDDTSTFYMCPP 67
QY 52 PVGSTVIRLE 61
DB 68 PSGSTLVRL 77

RESULT 10
US-08-042-747A-6
Sequence 6, Application US/08042747A
Patent No. 5487969
GENERAL INFORMATION:
APPLICANT: Eberle, Richard
APPLICANT: Black, Darla
APPLICANT: Scinicariello, Franco
APPLICANT: Hilliard, Julia K.
TITLE OF INVENTION: Cloning and Amplification of Monkey B
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cox & Smith Incorporated
STREET: 112 East Pecan Street, Suite 2000
CITY: San Antonio
STATE: Texas
COUNTRY: USA
ZIP: 78205
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042.747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Raymond, W. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395

TELEX: 767609
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 891 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-042-747A-6

Query Match 32.8%; Score 107; DB 1; Length 891;
Best Local Similarity 40.9%; Pred. No. 1.5e-05;
Matches 18; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 18 GTNSSPSTQNTVTSREVSVVOLSEESTFYLCPPVPGSTVIRLE 61
DB 72 GTNASVEAGHATLRENLRDIALDGDATFYVCPPTGATVVOFE 115

RESULT 11
US-08-220-151-4
Sequence 4, Application US/08220151
Patent No. 5529780
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Limbach, Keith J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
CANINE HERPESVIRUS gb, gc AND gd AND USES THEREFOR
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220.151
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-220-151-4

Query Match 32.5%; Score 106; DB 1; Length 1041;
Best Local Similarity 23.5%; Pred. No. 2.5e-05;
Matches 28; Conservative 14; Mismatches 17; Indels 60; Gaps 3;

QY 3 YFRNCIF-----FLVILY-----FLVILY-----17
DB 22 YFRQRCFFPSLIGIAATGSRHNGSSGLTRLARYVSFIWLVLFVGRPVESQSGTSEQ 81
QY 18 -----GTNSSPSTQNTVTSREVSVVOLSEESTFYLCPPVPGSTVIRLE 61
DB 82 PRTVATPEVGGTTPPKPTTDPDMSDMREALRASQIEANGPSTFYMCPPPSGSTVVRLE 140

RESULT 12
US-08-413-118-4
; Sequence 4, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-4

Query Match 32.5%; Score 106; DB 1; Length 1041;
Best Local Similarity 23.5%; Pred. No. 2.5e-05;
Matches 28; Conservative 14; Mismatches 17; Indels 60; Gaps 3;

QY 3 YFRRNCIF-----FLVILY----- 17
|||:| | : :|:
Db 22 YFQRCFFPSLLGIAATGSRHNGSSGLRLARYVSFIWILFLVGPVPVGGSGTSEQ 81
QY 18 -----GTNPSPTQNTVSREVSVSVQVLSSEE-----STFYLCPPPVGTVIRLE 61
Db 82 PRRTVATPEVGGTPPKPTDPTDMSDMDREALRASQIEANGPSTFYMCPPPSGTVVRLE 140

RESULT 13
US-08-473-446-4
; Sequence 4, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.

; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-473-446-4

Query Match 32.5%; Score 106; DB 3; Length 1041;
Best Local Similarity 23.5%; Pred. No. 2.5e-05;
Matches 28; Conservative 14; Mismatches 17; Indels 60; Gaps 3;

QY 3 YFRRNCIF-----FLVILY----- 17
|||:| | : :|:
Db 22 YFQRCFFPSLLGIAATGSRHNGSSGLRLARYVSFIWILFLVGPVPVGGSGTSEQ 81
QY 18 -----GTNPSPTQNTVSREVSVSVQVLSSEE-----STFYLCPPPVGTVIRLE 61
Db 82 PRRTVATPEVGGTPPKPTDPTDMSDMDREALRASQIEANGPSTFYMCPPPSGTVVRLE 140

RESULT 14
US-08-220-151-7
; Sequence 7, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435

Search completed: October 8, 2003, 17:09:34
Job time : 2.77075 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:10 ; Search time 5.11858 Seconds
(without alignments)
1953.624 Million cell updates/sec

Title: US-09-147-052-2_COPY_1_63

Perfect score: 326

Sequence: 1 MHYFRNCIFFLIVLYGYN.....STFYLCPPPGVSTVIRLERG 63

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	326	100.0	456 18 AAW36050	Hybrid Marek's dis
2	314	96.3	865 14 AAR30169	Marek's Disease VI
3	314	96.3	1086 18 AAW36051	Hybrid Marek's dis
4	107.5	33.0	879 16 AAR77024	Canine herpesvirus
5	107	32.8	891 17 AAR92746	B virus gB glycopr
6	107	32.8	891 19 AAW70293	Simian herpesvirus
7	100	30.7	854 8 AAP70347	Varicella-zoster v
8	100	30.7	868 13 AAR22615	Varicella-zoster v
9	99.5	30.5	250 22 AAW50114	Feline herpesvirus

10	99.5	30.5	943	21	AAV32470	DNA encoding feline
11	99.5	30.5	943	22	AAW50113	Feline herpesvirus
12	96	29.4	885	17	AAR92747	SAB virus gB glyco
13	95	29.1	400	8	AAP70654	Sequence encoded b
14	95	29.1	694	12	AAR14666	Truncated HSVgB po
15	95	29.1	903	7	AAP60244	Herpes simplex vir
16	95	29.1	903	8	AAP70426	Herpes simplex vir
17	95	29.1	903	8	AAP71135	Recombinant herpes
18	95	29.1	904	12	AAR14665	Herpes Simplex Vir
19	95	29.1	904	14	AAR14779	HSVgB polypeptide.
20	95	29.1	904	17	AAW00375	Glycoprotein B (gB
21	95	29.1	904	18	AAW34553	HSV-1 glycoprotein
22	95	29.1	904	22	AAW74441	Herpes simplex vir
23	95	29.1	905	9	AAP80915	Herpes simplex vir
24	95	29.1	973	12	AAR14680	Sequence of Herpes
25	87	26.7	973	8	AAP70769	Glycoprotein B of
26	85	26.1	854	19	AAW72113	HSV-2 strain SB5 C
27	85	26.1	904	9	AAP80914	Sequence of Herpes
28	85	26.1	904	14	AAR41778	Glycoprotein B (gB
29	85	26.1	904	17	AAW00376	HSV-2 glycoprotein
30	85	26.1	904	18	AAW34552	Herpes simplex vir
31	85	26.1	904	19	AAW72193	HSV-2 strain SB5 C
32	85	26.1	904	22	AAW74442	Herpes simplex vir
33	85	26.1	904	23	AAW17812	Herpes simplex vir
34	85	26.1	907	8	AAP711136	Herpes simplex vir
35	84	25.8	845	6	AAP50035	N-terminal sequenc
36	84	25.8	913	19	AAW68404	Aujeszky's disease
37	82.5	25.3	928	16	AAR77399	BHV1 gI glycoprote
38	82.5	25.3	932	19	AAW44947	Bovine herpesvirus
39	82.5	25.3	933	13	AAR27807	Bovine herpes viru
40	82.5	25.3	933	14	AAW41343	Bovine herpesvirus
41	74	22.7	903	6	AAP50312	Herpes simplex vir
42	64	19.6	1605	22	ABW70375	Drosophila melanog
43	61	18.7	61	23	ABP39693	Staphylococcus epi
44	60.5	18.6	164	24	ABU70485	Human adipocyte Se
45	60.5	18.6	444	18	AAW36845	Single chain T-cell

ALIGNMENTS

RESULT 1

AAW36050

ID AAW36050 standard; Protein; 456 AA.

XX AC

AAW36050;

XX DT

15-JUL-1998 (first entry)

XX DE

Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX DE

Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

KW KW

antigen; vaccine; poultry.

XX XX

Chimeric - Marek's disease gammaherpesvirus.

OS OS

Chimeric - Mycoplasma gallisepticum.

XX FH

Key Location/Qualifiers

FT Region

1..64

FT Region

/note= "derived from Marek's disease virus gB protein"

65..456

FT FT

/note= "derived from M. gallisepticum antigenic protein"

XX XX

WO9736924-A1.

PN PN

09-OCT-1997.

XX PD

28-MAR-1997;

97WO-JP01084.

XX PF

29-MAR-1996;

96JP-0103548.

XX PR

(JAPG) NIPPON ZEON KK.

PI Saito S, Tsuzaki Y, Yanagida N;
XX WPI: 1997-503046/46.
DR N-PSDB; AAT96595.
XX
PT Fusion protein comprising herpes virus outer membrane protein and
PT antigenic polypeptide - for prevention of infection by Mycoplasma
PT gallisepticum, especially in poultry
XX
PS Disclosure: Page 16-19; 51pp; Japanese.
XX
CC This sequence represents the chimeric protein 40 K-S which comprises a
CC fragment of the Marek's disease virus outer membrane protein gB fused
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
CC protein can be used in recombinant live vaccines for prevention of
CC infection by Mycoplasma gallisepticum, especially as the outer membrane
CC protein shows antigenicity in poultry.
XX
SQ Sequence 456 AA;
Query Match 100.0%; Score 326; DB 18; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.6e-33;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTVSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVLYGTNSPSTQNTVSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 EFG 63
DB 61 EFG 63
RESULT 2
AAR30169
ID AAR30169 standard; Protein; 865 AA.
XX
AC AAR30169;
XX
DT 25-MAR-2003 (updated)
DT 07-MAY-1993 (first entry)
XX
DE Marek's Disease Virus glycoprotein B homologue of HSV.
XX
KW Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
KW Herpes Simplex Virus.
XX
OS Marek's Disease Virus.
XX
PN EP520753-A1.
XX
PD 30-DEC-1992.
XX
PF 24-JUN-1992; 92EP-0305775.
XX
PR 28-JUN-1991; 91US-0722860.
PR 10-DEC-1991; 91US-0803633.
XX
PA (JAPG) NIPPON ZEON KK.
PA (USDA) US SEC OF AGRIC.
XX
XX Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
XX WPI: 1993-001546/01.
XX
DR Recombinant fowl pox virus contg. Marek's disease virus antigen
PT gene - used to produce cell-free vaccine against Marek's disease
PT virus
XX
PS Example 2; Page 15-19; 30pp; English.
XX
CC The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
CC cf. MDV GA strain was cloned into pUC18. A 2.8kb BamHI-Sali
CC

CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
CC from K3 fragment were ligated with EcoRI, Sali digested pUC18. The
CC sequence of the putative MDV gBh was determined by sequencing a set
CC of deletion mutants. The nucleotide and amino acid sequences were
CC found to be identical with the published sequences of the gBh of
CC RBIB strain of MDV (Ross et al., J. Gen. Virol., 70:1789-1894, 1988).
CC A fragment contg. the entire coding region of MDV gBh was inserted
CC into pNZ1729R (see AAQ34774-Q34778) to produce a recombinant FPV/MDVgBh
CC virus for immunising chickens. See also AAQ34780-Q34781.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 865 AA;
Query Match 96.3%; Score 314; DB 14; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.9e-31;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVLYGTNSPSTQNTVSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
DB 1 MHYFRNCIFFLIVLYGTNSPSTQNTVSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
QY 61 E 61
DB 61 E 61
RESULT 3
AAW36051
ID AAW36051 standard; Protein; 1086 AA.
XX
AC AAW36051;
XX
DT 15-JUL-1998 (first entry)
XX
DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX
KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW antigen; vaccine; poultry.
XX
OS Chimeric - Marek's disease gammaherpesvirus.
OS Chimeric - Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Region 1..672
FT /note= "derived from Marek's disease virus gB protein"
FT Region 693..1086
FT /note= "derived from M. gallisepticum antigen"
XX
PN W09736924-A1.
XX
PD 09-OCT-1997.
XX
PF 28-MAR-1997; 97WO-JP01084.
XX
PR 29-MAR-1996; 96JP-0103548.
XX
PA (JAPG) NIPPON ZEON KK.
XX
PI Saito S, Tsuzaki Y, Yanagida N;
XX WPI: 1997-503046/46.
XX N-PSDB; AAT96596.
XX
PS Fusion protein comprising herpes virus outer membrane protein and
XX antigenic polypeptide - for prevention of infection by Mycoplasma
XX gallisepticum, especially in poultry
XX
PS Disclosure: Page 22-30; 51pp; Japanese.
XX
CC This sequence represents the chimeric protein 40 K-C which comprises a
CC fragment of the Marek's disease virus outer membrane protein gB fused
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
CC protein can be used in recombinant live vaccines for prevention of

CC	infection by Mycoplasma gallisepticum, especially as the outer membrane
CC	protein shows antigenicity in poultry.
XX	
SQ	Sequence 1086 AA;
	Query Match 96.3%; Score 314; DB 18; Length 1086;
	Best Local Similarity 100.0%; Pred. NO. 6.6e-31;
	Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MHYFRNCIFFLVILYGTNSSPSTQNTSVREVSVSSVOLSEESTFVLCPPPVGSGTVRL 60
Db	1 MHYFRNCIFFLVILYGTNSSPSTQNTSVREVSVSSVOLSEESTFVLCPPPVGSGTVRL 60
QY	61 E 61
	-
Db	61 E 61
RESULT 4	
AAR77024	
ID	AAR77024 standard; Protein; 879 AA.
XX	
AC	AAR77024;
XX	
DT	20-APR-1996 (first entry)
XX	
DE	Canine herpesvirus gB homologue.
XX	
DE	
XX	
KW	CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;
KW	canarypox virus; ALVAC; VCP320; dog; puppy; immunisation; antigen;
KW	vaccine.
XX	
XX	Canine herpesvirus.
OS	
XX	
FH	Key Location/Qualifiers
FT	Region 725..741
FT	/label= Transmembrane_region
FT	Region 747..771
FT	/label= Transmembrane_region
XX	
XX	WQ9526751-A1.
PN	
XX	
PD	12-OCT-1995.
XX	
PF	30-MAR-1995; 95WO-US03982.
XX	
PR	29-MAR-1995; 95US-0413118.
XX	
PR	30-MAR-1994; 94US-0220151.
XX	
PA	(VIRO-) VIROGENETICS CORP.
XX	
PI	Limbach KJ, Paoletti E;
XX	
DR	WPI; 1995-366131/47.
XX	
DR	N-PSDB; AAT01402.
XX	
PT	Nucleic acids encoding canine herpes virus (CHV) gB, gC and gD
PT	glyco:proteins - also glyco:proteins and vectors, for the
PT	immunisation of neonatal puppies and adult dogs against CHV
XX	
PS	Example 3; Fig 1A-G; 241pp; English.
XX	
CC	A canine herpesvirus (CHV) protein has a predicted amino acid
CC	sequence (AAR77024) that shows significant homology with the gB
CC	glycoprotein of numerous herpesviruses. It is the product of a
CC	gene (see AAT01402) isolated from CHV genomic DNA using a probe contg.
CC	the feline herpesvirus gB, gC and gD genes. CGV glycoprotein gC and
CC	gD homologues (AAR77025-26) have also been obt'd. These glycoproteins,
CC	including recombinant glycoproteins expressed from attenuated
CC	recombinant virus vectors, e.g. ALVAC recombinant vcp320 (see
CC	AAT01406), can be used in antigenic, immunological or vaccine
CC	compositions to protect puppies and adult dogs against CHV.
XX	

```

Db      72 GTNASVEAGHATLRENLRIKALDGDATFFVCPPTGATVVQFE 115

RESULT 6
AAW70293
ID      AAW70293 standard; Protein; 891 AA.
XX
XX      AAW70293;
XX
XX      06-NOV-1998 (first entry)
XX
XX      Simian herpesvirus B gB glycoprotein sequence (UL27).
XX
XX      Simian herpesvirus B gB glycoprotein; UL27; ICP protein; UL28;
KW      differential diagnostic test; immunoassay; antibody.
XX
XX      Simian herpesvirus B.
XX
XX      US5767265-A.
XX
XX      16-JUN-1998.
XX
XX      10-OCT-1995; 95US-0541878.
XX
XX      01-APR-1993; 93US-0042747.
XX      10-OCT-1995; 95US-0541878.
XX
XX      (SWBI-) SOUTHWEST FOUND.BIOMEDICAL RES.
XX
XX      Black D, Eberle R, Hilliard J, Scinicariello F;
XX      WPI; 1998-361791/31.
XX      N-PSDB; AAV33167.
XX
XX      Monkey herpes B virus DNA - coding for gB glycoproteins and
XX      polypeptides
XX
XX      Claim 2; Columns 19-24; 22pp; English.
XX
XX      The invention provides a Simian herpesvirus B DNA sequence coding
XX      for the present gB glycoprotein (UL27) and a portion of an ICP
XX      18.5 kDa protein (UL28; AAW70294). The invention uses these DNA and
XX      protein sequences as a basis for the development of differential
XX      diagnostic tests for the rapid identification of Simian herpesvirus B
XX      cases. Therefore, the virus can be detected by detecting the DNA
XX      sequence and knowledge of the amino acid sequence will help in the
XX      design of DNA probes and of peptides for use in immunoassays and for
XX      antibody production.
XX
XX      Sequence 891 AA;
XX
XX      Query Match 32.8%; Score 107; DB 19; Length 891;
XX      Best Local Similarity 40.9%; Pred. No. 0.00013;
XX      Matches 18; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
XX
Qy      18 GTNSSPSTQNTREVSSVQLSSESTFYICPPPGVSTVIRLE 61
       III:I : I I I : : : : III:III I I I : I
Db      72 GTNASVEAGHATLRENLRIKALDGDATFFVCPPTGATVVQFE 115

RESULT 7
AAP70347
ID      AAP70347 standard; Protein; 854 AA.
XX
XX      AAP70347;
XX
XX      25-MAR-2003 (updated)
XX      03-OCT-2002 (updated)
XX      18-MAY-1991 (first entry)
XX
XX      Varicella-zoster virus (VSV) immunogenic outer surface protein.
XX
XX      'Varicella-zoster virus; gB gene product; varicella; vaccine.

```

```

XX PT Vaccine against varicella-zoster virus (VZV) - produced from gene
XX PT encoding outer surface viral protein
XX PS Claim 1; Page 9; 18pp; English.
XX CC Chickenpox is caused by varicella-zoster virus (VZV), a member of
XX CC the herpesvirus gp. VZV has five major glycoproteins on its surface
XX CC which are the prods. of three different genes, gA, gB and gC.
XX CC Monoclonal antibodies to gA and gB display complement-independent
XX CC neutralisation and the monoclonal antibodies of gC display
XX CC complement-dependent neutralisation. Cytoplasmic RNAs were prepd.
XX CC from VZV-infected MRC-5 cells. The RNAs encoded by the different
XX CC VZV HindIII fragments were selected by hybridisation to cloned VZV
XX CC HindIII DNA fragments (J.R.Ecker & R.W. Hyman, Proc. Natl. Acad.
XX CC Sci. USA 79:156 (1982)) bound to nitrocellulose. These RNAs were
XX CC were immunopptd. by polyclonal monospecific guinea pig antibodies
XX CC raised to gB purified by monoclonal antibody affinity
XX CC chromatography. By this analysis it was found that a 100 kD in
XX CC vitro translation prod. from mRNA selected by the VZV-HindIII-D
XX CC fragment could be immunopptd. by the anti-gB antibodies which
XX CC neutralise viral infectivity. The polypeptide may be used to
XX CC react with human convalescent zoster sera and with monospecific
XX CC antisera which neutralise viral infectivity. The polypeptide or
XX CC fragments of it are useful for the prepn. of vaccines or diagnosis
XX CC of VZV infection.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX PS Sequence 868 AA;
XX PS Query Match 30.7%; Score 100; DB 13; Length 868;
XX PS Best Local Similarity 54.8%; Pred. No. 0.00098;
XX PS Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Oy 31 REVSSVQLSEESTFYLCPPPGVSTVIRLE 61
Db 41 REAIHKSDAETRTFFVCPPTGTVIRLE 71

RESULT 9
AAM50114
XX ID AAM50114 standard; Protein; 250 AA.
XX AC AAM50114;
XX DT 21-DEC-2001 (first entry)
XX DE Feline herpesvirus glycoprotein B recombinant antigen PFHVgB250.
XX KW FHV; glycoprotein B; PFHVgB250; antigen; immune status;
XX KW vaccination status; cat.
XX OS Feline herpesvirus.
XX XN W0200166568-A2.
XX PD 13-SEP-2001.
XX PF 07-MAR-2001; 2001WO-US07251.
XX PR 09-MAR-2000; 2000US-0521738.
XX XX (HESK-) HESKA CORP.
XX PA (COLS ) UNIV COLORADO STATE RES FOUND.
XX PI Jensen WA, Lappin MR, Rosen DK, Andrews JS;
XX XX WPI; 2001-639000/73.
XX DR N-PSDB; AAH27061.
XX XX Determining immune status or vaccination status of an animal to e.g.
XX PT calicivirus comprises using a recombinant viral antigen -

```

```

XX PS Claim 4; Page 91-92; 132pp; English.
XX CC The present sequence is that of truncated feline herpesvirus
XX CC glycoprotein B recombinant antigen PFHVgB250. The recombinant
XX CC antigen, fused to an N-terminal His tag, was produced in Escherichia
XX CC coli cells transformed by a recombinant vector comprising nucleic
XX CC acid nFHVgB750 (see AAH27061). PFHVgB250 is an example of a
XX CC recombinant infectious agent antigen that can be used in the method
XX CC of the invention to determine the immune status of an animal.
XX CC The method involves contacting a biological specimen of an animal
XX CC (cat, dog or horse) with a recombinant antigen, and detecting the
XX CC presence or absence of a complex between the recombinant antigen
XX CC and an antibody present in the sample. The method determines
XX CC whether the animal is protected against disease or should be
XX CC vaccinated. Recombinant antigens (see AAM50107-24), nucleic acids
XX CC encoding them (see AAH27054-71), methods of producing them, and
XX CC assay methods are provided.
XX PS Sequence 250 AA;
XX PS Query Match 30.5%; Score 99.5; DB 22; Length 250;
XX PS Best Local Similarity 24.6%; Pred. No. 0.00024;
XX PS Matches 29; Conservative 14; Mismatches 16; Indels 59; Gaps 4;

Oy 3 YFRNCIF-----FLVILY-----GTNssp 23
Db 22 YFQRCFFPSILGTAATGSRHNGSSGLTRLARYVSFIWLVFLVGPVPBGQSGSTSEQ 81
Oy 24 STQNTVSREV-----VSSVQLSEEE-----STFYLCPPPGVSTVIRLE 61
Db 82 PRRTVATPEVGVHHQNLQIPPICRYEEALRASQIEANGPSTFYMCPPPGSTVIRLE 139

RESULT 10
AAY32470
XX ID AAY32470 standard; Protein; 943 AA.
XX AC AAY32470;
XX DT 27-MAR-2000 (first entry)
XX DE DNA encoding feline herpesvirus-1 glycoprotein B.
XX KW FHV; glycoprotein B; raccoon poxvirus; vaccine; cat;
XX KW feline viral rhinotracheitis.
XX OS Feline herpesvirus type 1.
XX PH Key Location/Qualifiers
XX FT Peptide 1..66
XX FT Domain /note= "signal peptide"
XX FT Domain /note= "extracellular domain"
XX FT Domain /note= "transmembrane domain"
XX FT Domain /note= "cytoplasmic domain"
XX FT Modified-site 152..154 /note= "Asn is N-glycosylated"
XX FT Modified-site 261..263 /note= "Asn is N-glycosylated"
XX FT Modified-site 364..366 /note= "Asn is N-glycosylated"
XX FT Modified-site 405..408 /note= "Asn is N-glycosylated"
XX FT Modified-site 526..528 /note= "Asn is N-glycosylated"
XX FT Modified-site 610..612 /note= "Asn is N-glycosylated"
XX FT Modified-site 643..645 /note= "Asn is N-glycosylated"
XX FT Modified-site /note= "Asn is N-glycosylated"
XX

```


WPI; 1996-105220/11.
N-PSDB; AAT16480.

Detection of herpes B virus by PCR amplification of sample DNA - to detect a specific herpes simian monkey B virus DNA segment.

Disclosure; Column 29-34; 22pp; English.

This sequence encodes the herpes simian monkey SA8 virus gB glycoprotein. Proteins such as the gB glycoprotein have potential use in the development of serological immunoassays. One approach is to synthesize peptides which, based on the properties of the predicted protein sequence, are likely to be immunologically active. Such peptides can be used as substrate antigens in immunoassays to detect serum antibodies which recognize this specific peptide sequence. Synthetic peptides may also be used to produce antibodies against specific regions of the gB glycoprotein which are unique to one virus. These can then be used to develop virus-specific immunoassays for differentiation of SA8 virus from other primate alpha-herpes viruses and for identification of antibodies directed against SA8 virus in primate serum samples.

Query Match 29.4%; Score 96; DB 17; Length 885;
Best Local Similarity 38.1%; Pred. No. 0.0032;
Matches 16; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 20 NSSPSTQNTVTSREWSSVQLSEESTYLCPPPVGVSTVIRLE 61
I:I | | | : : : : : | | | | | : | : | : |
DB 68 NASVEAGRTLREDLRIKARDGDATFYVCPPTGATVWQFE 109

RESULT 13
AAP70654
ID AAP70654 standard; Protein; 400 AA.
XX AAP70654;
XX
XX 25-MAR-2003 (updated)
DT 16-APR-1991 (first entry)
XX
XX Sequence encoded by g B gene of Herpes simplex virus (HSVgB).
XX Yeast expression vector.
XX Herpes simplex virus.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..30
FT /label= Signal peptide
FT Protein 31..400
XX
XX EP216195-A.
PN
XX
XX 01-APR-1987.
PD
XX 29-AUG-1986; 86EP-0112005.
PF
XX 30-AUG-1985; 85JP-0192916.
PR
XX (KAGA) CHEMO SERO THERAPEUTIC RES INS.
XX
XX Nozaki C, Nakatake H, Kino Y, Eto T, Makizumi K, Ohtomo N;
PI
XX WPI; 1987-087941/13.
DR N-PSDB; AAN70101.
XX
XX Plasmid contg. herpes simplex virus gene - used for transforming
PT yeast for producing HSV vaccines and diagnostic reagents
XX
XX Disclosure; Fig 1; 38pp; English.
XX

```

XX AAP60244;
AC
XX
XX 25-MAR-2003 (updated)
DT 31-JUL-1991 (first entry)
XX
DE Herpes simplex virus glycoprotein gB.
DE
XX HSV; gb glycoprotein; vaccine.
XX
XX Herpes simplex virus.
OS
XX EP170169-A.
PN
XX 05-FEB-1986.
PD
XX
XX 19-JUL-1985; 85EP-0109042.
PF
XX
XX 20-JUL-1984; 84JP-0151766.
PR
XX 11-DEC-1984; 84JP-0262465.
PR
XX (KAGA ) CHERO SERO THERAPEUTIC RES INST.
PA
XX
XX Nozaki C, Makizumi K, Kino Y, Eto T, Ohtomo N;
PI
XX
XX WPI: 1986-036935/06.
DR
XX N-PSDB; AAN60195.
DR
XX
PT Recombinant DNA containing herpes simplex virus gene or fragment -
PT useful in transformant yeast for prodn. of high purity herpes
PT simplex virus glycoproteins for use in protective vaccines.
XX
XX Disclosure; Fig 6; 53pp; English.
PS
XX
XX The gB glycoprotein is useful for the production of vaccines
CC conferring protection against herpes simplex virus infections.
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC
XX
XX Sequence 903 AA;
SQ
Query Match 29.1%; Score 95; DB 7; Length 903;
Best Local Similarity 34.1%; Pred. No. 0.0045;
Matches 15; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
QY 18 GTNSPSTQNTVTSRETVSSVOLSEESTFYLCPPPVGSTVIRLE 61
DB 84 GDNATVAAGHATLREHLRDKAENTDANFYVCPPTGATVQFE 127

```

Search completed: October 8, 2003, 17:02:48
Job time : 7.11858 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:58:45 ; Search time 39.2516 Seconds
(without alignments)
2997.887 Million cell updates/sec

Title: US-09-147-052-2
Perfect score: 2324
Sequence: 1 MHYFRNCIFLVLIVLYCTN.....SSNENNAKIPGRRPGTFL 456

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTREMBL_23.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organelle.*
- 10: sp_phage.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_virus.*
- 14: sp_vertebrate.*
- 15: sp_unclassified.*
- 16: sp_rvirus.*
- 17: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200	51.6	271	2 Q49464	Q49464 mycoplasma
2	845	36.4	671	2 Q9L8D6	Q9L8D6 mycoplasma
3	832	35.8	702	2 Q49499	Q49499 mycoplasma
4	829	35.7	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	34.6	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	32.7	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	32.4	702	2 Q9498	Q49498 mycoplasma
8	750.5	32.3	645	2 Q92ID1	Q92id1 mycoplasma
9	749.5	32.3	644	2 Q92HR9	Q92hr9 mycoplasma
10	749	32.2	650	2 Q49495	Q49495 mycoplasma
11	745.5	32.1	649	2 Q49497	Q49497 mycoplasma
12	740.5	31.9	486	2 Q8RLX9	Q8rlx9 mycoplasma
13	715.5	30.8	647	2 Q49468	Q49468 mycoplasma
14	704	30.3	656	2 Q9KH15	Q9kh15 mycoplasma
15	700.5	30.1	703	2 Q05122	Q05122 mycoplasma
16	619.5	26.7	419	2 Q9L8D5	Q9L8D5 mycoplasma

17	578	24.9	386	2 Q49500	Q49500 mycoplasma
18	557.5	24.0	367	2 Q9XCG7	Q9xcg7 mycoplasma
19	314	13.5	805	12 Q98Y44	Q98y44 turkey herp
20	314	13.5	805	12 Q98Y45	Q98y45 turkey herp
21	314	13.5	865	12 Q83291	Q83291 marek disea
22	314	13.5	865	12 Q8JLW2	Q8jlw2 turkey herp
23	314	13.5	865	12 Q8JLW3	Q8jlw3 turkey herp
24	314	13.5	865	12 Q8JLW4	Q8jlw4 turkey herp
25	196.5	8.5	864	12 Q69408	Q69408 meleagridd h
26	196.5	8.5	870	12 Q9DPQ9	Q9dpq9 meleagridd h
27	196.5	8.5	870	12 Q9E1G4	Q9e1g4 meleagridd h
28	189.5	8.2	320	2 Q49496	Q49496 mycoplasma
29	185.5	8.0	9439	16 Q8CP76	Q8cp76 staphylococ
30	177.5	7.6	865	12 Q89406	Q89406 turkey herp
31	177.5	7.6	865	12 Q9PWZ1	Q9p wz1 turkey herp
32	177	7.6	6713	16 Q99U54	Q99u54 staphylococ
33	177	7.6	6713	16 Q931R6	Q931r6 staphylococ
34	173.5	7.5	933	2 Q53653	Q53653 staphylococ
35	173	7.4	946	16 Q8NXJ1	Q8nxj1 staphylococ
36	173	7.4	1302	2 Q49547	Q49547 mycoplasma
37	170	7.3	9904	16 Q8NWQ6	Q8nwq6 staphylococ
38	168.5	7.3	810	5 Q814Y6	Q814y6 plasmodium
39	166.5	7.2	3890	16 Q99U53	Q99u53 staphylococ
40	166.5	7.2	4688	16 Q9PQ08	Q9pq08 ureaplasma
41	166	7.1	661	16 Q8YMJ8	Q8ymj8 anabaena sp
42	165.5	7.1	2481	16 Q99QR6	Q99qr6 staphylococ
43	164.5	7.1	1774	5 Q81AU8	Q81au8 plasmodium
44	163.5	7.0	2462	16 Q8RGZ3	Q8rgz3 fusobacteri
45	163.5	7.0	2806	16 Q8RI19	Q8ri19 fusobacteri

ALIGNMENTS

RESULT 1

Q49464	PRELIMINARY;	PRT;	271 AA.
ID Q49464			
AC Q49464;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE TM-1 (Fragment).			
GN TM-1.			
OS Mycoplasma gallisepticum.			
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX NCBI_TaxID=2096;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=94025893; PubMed=8212828;			
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,			
RA Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.;			
RT "Cloning and DNA sequence of a 29 kilodalton polypeptide gene of			
RT Mycoplasma gallisepticum as a possible protective antigen.";			
RL Vaccine 11:1061-1066(1993).			
DR EMBL; S65869; AAB28343.2; .			
FT NON_FER	271	271	
SQ SEQUENCE	271 AA;	29817 MW;	8B25DE0CD5C85CA2 CRC64;

Query Match	51.6%;	Score 1200;	DB 2;	Length 271;
Best Local Similarity	97.1%;	Pred. No. 1.5e-48;		
Matches 238;	Conservative 3;	Mismatches 4;	Indels 0;	Gaps 0;
QY	64	CMSITKDKANPNNGQTQLEARNMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN	123	
Db	27	CMSITKDKANPNNGQTQLEARNMELTDLINAKAMTILASLDYAKIEASLSAYSEATVN	86	
QY	124	NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEQRTNLEGLS	183	
Db	87	NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEQRTNLEGLA	146	
QY	184	STAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTKNNKNTLSTINQKTN	243	
Db	147	STAYNOIRNNLVLYNKASSLITKTDPLNGGTLDSNEITTKNNKNTLSTINQKTN	206	

QY 244 ADALSNFIKKVIONNEOSFGVTETNANVQPSNTSFVAFSADVTPTVNTKYARRTVWNGDE 303
 DB 207 ADALSNFIKKVIONNEOSFGVTETNANVQPSNTSFVAFSADVTPTVNTKYARRTVWNGDE 266
 QY 304 PSSRI 308
 DB 267 PSSRI 271

RESULT 2
 Q9L8D6
 ID Q9L8D6 PRELIMINARY; PRT; 671 AA.
 AC Q9L8D6;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE PMGA-like protein 9.2.
 OS Mycoplasma gallisepticum.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F;
 RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
 RA May J.D., Hughlett M.B.; from the F-strain (vaccine strain) of
 RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
 RT Mycoplasma gallisepticum."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF210770; AAF29524.1; -.
 SQ SEQUENCE 671 AA; 71898 MW; 4D6A56B59175D679 CRC64;

Query Match 36.48; Score 845; DB 2; Length 671;
 Best Local Similarity 41.28; Pred. No. 1.1e-31;
 Matches 186; Conservative 77; Mismatches 127; Indels 62; Gaps 10;
 QY 53 VGSTVIRLEFGCMSITTKDANP-----NGQT-----OLEAA 84
 DB 15 IGSFVMLAAASCTSATPTPEPKPDMPNPPSGDNGGDTNPGDGGMENSAAQQLAA 74
 QY 85 RMEITDLINAKAMTASLQDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAL 144
 DB 75 KKEISDLATQSSNLAKYADYNTIONTLTAAYTTAKSTSDNTSVTLQVKSATSTLQAAI 134
 QY 145 NQANTDKTDENEHPLVEAYKALKTTLEQATNLEGLSTAYNQIRNNLVLYNKASSL 204
 DB 135 DTAASSKTSFDEKNPELIKAYALKETLKNEETVLSGLTDSNFATIKTNLTALYQSGKF 194
 QY 205 ITRTLDPLNGTLLDSNEITANKNNITLSTNEQKTNADALSNSFIKKVIONNEOSFV 264
 DB 195 VKATLDPVSGNA-POADIITKADKDADIADAVSKLETWNTANTLATSFKVEVLKNTLGI 253
 QY 265 GTTFNANVQPSNTSFVAFSADVT-----VNYKYARRTVWNGDEPSSRILA-----NTNSIT 316
 DB 254 DT-TNNREOPGNTSFVGSYSVNATNNNEIPNWNFAQRKVWTSNDRGTSLSISSTL 312
 QY 317 DVSWIYSLAGTNTKYQSFSGNYGPGCYLYFPKLYKAADANNVGLQKLNNGVQVQVEF 376
 DB 313 EVSWIYSLAGTNTKYQSFSGNYGPGCYLYFPKLYKAADANNVGLQKLNNGVQVQVEF 372
 QY 377 A-----TSTSA-----NNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTIE 414
 DB 373 APVTKTSVSDSGSDSNQNTESAETMPVTSDLPNAPTVSDINAKTLNLSNLFSGNTIE 432
 QY 415 LSVPTGEGNKNKVPAMIGNIYLLSSNENNAADKI 446
 DB 433 FSVPTGPS--NKVAPMIGNIYLLSNIANEAKV 462

RESULT 3
 Q49499
 ID Q49499 PRELIMINARY; PRT; 702 AA.
 AC Q49499;

DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PMGA1.4 protein precursor.
 GN PMGA1.4.
 OS Mycoplasma gallisepticum.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S6;
 RA MEDLINE=95010739; PubMed=7925999;
 RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
 RA Browning G.F., Whithear K.G., Walker I.D.;
 RT "The organisation of the multigene family which encodes the major cell
 RT surface protein, PMGA, of Mycoplasma gallisepticum."
 RL FEBS Lett. 352:347-352(1994).
 DR EMBL; L28424; AAA62418.1; -.
 KW Signal.
 FT SIGNAL
 SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;

Query Match 35.88; Score 832; DB 2; Length 702;
 Best Local Similarity 40.98; Pred. No. 4.5e-31;
 Matches 177; Conservative 86; Mismatches 124; Indels 46; Gaps 9;
 QY 50 PPPVGSTVIRLEFGCMSITTKDANPNG-----QTOLEAARMELTDLINAKAMTASLQ 103
 DB 54 PNPGGMMGMMNGG-----NTNPGNGGTDNAAQAAAKKESLDLLATONLSIYA 106
 QY 104 DYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLVE 163
 DB 107 DYANIONTLTAAYTTAKSTSDNTSATLEQVKSATSTLQTAIDTAASSKTSFDEKNPELIK 166
 QY 164 AYKALKTTLEQATNLEGLSTAYNQIRNNLVLYNKASSLITKLDPLNGCTLLDSNEI 223
 DB 167 AYNALKETLKWRNLSGLTDSNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAV 225
 QY 224 TTANKNNITLSTNEQKTNADALSNSFIKKVIONNEOSFVGTETNANVQPSNTSFVAFS 283
 DB 226 SQANTNISNAVSKLETWNTANTLATSFKVEVLKNTLGTIDT-TNNQEQPGNYSFVGS 284
 QY 284 ADVTP-----VNYKYARRTVWNGD-----EPSSRIANTNSITDVSWIYSLAGTNTKY 332
 DB 285 VDVTTGSDNARPNMSFAQRKVWTSNTDILSQPQAEGENQOSAPDVSWIYNTLTCMGAKYS 344
 QY 333 FFSFNYGPGCYLYFPKLYKAADANNVGLQKLNNGVQVQVEFATS-----TSAN 383
 DB 345 LTFNYGPGCYLYFPKLYKAADANNVGLQKLNNGVQVQVEFATS-----TSAN 383
 QY 384 N-----TTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNKNKVPAMIG 432
 DB 405 NRSTAAPAQSGSTEINPAPTLDIDIKIAKVTLSNLFKSGNTIEFSVPTAKGTSKVPAMIG 464
 QY 433 NYLSSNENNAADK 445
 DB 465 NMVLTSSDRDVK 477

RESULT 4
 Q9XCG8
 ID Q9XCG8 PRELIMINARY; PRT; 632 AA.
 AC Q9XCG8;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DE VlhA1 precursor (Fragment).
 GN VlhA1.
 OS Mycoplasma imitans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=29560;
 RN [1]
 RP SEQUENCE FROM N.A.

SQL	SEQUENCE	584 AA;	62453 MW;	5C467BA55FB27A72 CRC64;
	Query Match	34.6%;	Score 805;	DB 2; Length 584;
	Best Local Similarity	42.4%;	Pred. No. 6.5e-30;	
	Matches 180;	Conservative	Mismatches 113;	Indels 60; Gaps 10;
QY	73 NPNGQT----	OLEAARMELTDLINAKAMT	LASLDQYAKIEASLSAYSEATVNNLNA	128
Db	69 NPSGNTTPEOQLAAARKTL	TDLTGENTNVALVDYAKIQSTLSTAYMTAKTASENTSA	128	
QY	129 TLBOLKMAKNLESAINQANTKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYN	188		
Db	129 TLDNLRASFTLQAADKAASNKRTPDSANOPLVTAYNQLTKTLQSXTTSLEGLSENKY	188		
QY	189 QIRNNLVLYNKASSLTTKTLPUNGOTGLDSDNEITANKINNNTLSTINEOKTNADALS	248		
Db	189 SIKNHLSKLFDAGSAIAARTLDP--TWGTVEVMSVTKANEDIWTAWSKLTETWKTNADKP-	246		
QY	249 NSFIKKYVIONNESQSFVGTFTNAN--VOPSNYSFVAFSADVTP-----VNYYKARRT	297		
Db	247 NDPEKKPL--SKELVSTNDRAHNQEOPANWSFAGYSVDLTGTGSTGNSQLPNWNFAQR	304		
QY	298 VWGCD----EPSSRLANTNSITDVSWIYSLAGNTKYPQSFNSYGPSTGYLYPPYKLVKA	354		
Db	305 VWTSEGQOTGKTALVSPVSATIDVSWIYSLAGECTKYLTSFEYGGPTAFELYPPYKLVQ	364		
QY	355 ADANNVGLOYKLANGVQOVEF-----ATST-----S	381		
Db	365 ADSSVALQYLNKTSKLIINFPAKTMPTNDAQSENGVAVTTTEGRSSSEVLVADEVA	424		
QY	382 ANNTTAPTPAVDEIKVAKIVLSGLRGFGONTIELSVPTGBGNMKNVPIMIGNIYLSNEN	441		
Db	425 AVANEMNPPTVSDINIAKVTLSGLTGTENTIEFSVPT-----NKVAPMIGNMYLTNSNG	479		
QY	442 NADKI 446			
Db	480 SQSKI 484			
RESULT	6			
Q9KH14				
ID	Q9KH14	PRELIMINARY;	PRT;	680 AA.
AC	Q9KH14;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DF	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)			
DE	Adhesin pmgA1.3.			
NCBI	pmgA1.3.			

[illegible]

```

QY 184 STAYNOIRNLDVLYNKASSLITKTLDPNGGTLDSNEITANKNINNTL-----ST 236
Db 183 QEKYSAILSEINAASABEIKQTLNPNVG-----NLPVVAALNAENTKILEAIKEK 236
QY 237 INEQKTADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARR 296
Db 237 INSEKSNADLFANYQLYKL--DRTKLMSEGSNNTKQPGNYSFVAYASDIAPNWNFAOR 293
QY 297 TVWNGD-----EPSSRILANTNSITDVSWIYSLAGTNTTKYQFSNYPSTGYLYPPYKL 351
Db 294 TWTADTSRTWTSPLPNLQNSAPLTDVSWIYTLSGTGAKYTLTDFYGPOTGYLYPPYKL 353
QY 352 VRAADANNVGLQYKLNNGVQVEF-----ATSTSANNNTANTPTAV 393
Db 354 VTSD--KVGLOTKLQAQDPVIAQSEAAATASAPAEOTDGRSEATATANEKVNPMPSV 411
QY 394 DEIKVAKIVLSGLRFQONTIELSVPTGEGNMKNKVPMPMIGNIYLSNNENNAK 445
Db 412 NTINAKVTLNLSKFGSNIETFSVPMQDQNMKNKVPMPMIGNIYLSNNENNAK 463

RESULT 7
Q49498
ID Q49498 PRELIMINARY; PRT: 702 AA.
AC Q49498;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE PMGAI.3 protein precursor. 19, Last annotation update)
GN PMGAI.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RT Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62417.1; -.
FT SIGNAL.
KW SIGNAL.
SQ SEQUENCE 702 AA; 75537 MW; 27368915FEE57B9F CRC64;

Query Match 32.4%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 1.8e-27;
Matches 176; Conservative 62; Mismatches 127; Indels 59; Gaps 11;

QY 73 NPNNGQT-----OLEARMELTDLINAKAMWLASLODYAKIEASLSAYSEATVNNLNA 128
Db 69 NPNNGTPEQOLAARKTLTDLGTENTNVALYADYAKIQSTLTAYMTAKTASNTSA 128
QY 129 TLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEORATNLEGLSSTAYN 188
Db 129 TLENLSASTTLOAIDAANDKRVDSVQNPVLAAYNNLTKTKSTTSLGSENKYG 188
QY 189 QTRNNLDVLYNKASSLITKTLDPNG--GTLLDSNEITTANKNINNTLS--TINEQKNA 244
Db 189 GIKNLSKLFDTGSAITAKTLDTPTSGERTPLEKVE---ANNKIMASPESLKKWKGNA 245
QY 245 DALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARRVW 299
Db 246 DKF-NEFEKNPLSKELKASTDSTAHNQEQPANWSEAFYASVLDTSNQNLPNNFAQRKW 304
QY 300 NGD--EPSSRILANTN-STTDVSWIYSLAGTNTTKYQFSNYPSTGYLYPPYKLVKAA 356
Db 305 TSENQOPGKTALVSPSVATDVSWIYSLAGETKTLTTFEYGPONAFILYPLKLVKAA 364
QY 357 ANNVGLQYKLNNGVQVEF-----ATST-----ATST-----SAN 383
::: ||| ||| : : : |||

```

```

Db 365 SSSVALQYSLNKTSSKLINFKPAETVSTNTDOSENAVATTSTTEARSSYKVLVADEAATS 424
QY 384 NTTANTPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMKNKVPMPMIGNIYLSNNENNA 443
Db 425 NNEMNHTPIVSDINIAKIVTLGSLTGFGENITFESVPEG-----KVAPMIGNIYLSNSESQ 479
QY 444 DKX 446
Db 480 VKI 482

RESULT 8
Q9ZID1
ID Q9ZID1 PRELIMINARY; PRT: 645 AA.
AC Q9ZID1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE M9 protein.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family.";
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;

Query Match 32.3%; Score 750.5; DB 2; Length 645;
Best Local Similarity 40.6%; Pred. No. 2.4e-27;
Matches 177; Conservative 67; Mismatches 133; Indels 59; Gaps 12;

QY 53 VGSTVIRLEFGCMSITKK-----DANPNNGQ-----TOLEARMELTDLINA 94
Db 15 IGSFVMAAASCTTTPNPNPSSGGMNGDTPGDCQGMMAASQELAARMGLTIVFDS 74
QY 95 KAMTILASQDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTDKTTF 154
Db 75 KAKNLGLYVDYKKTQDTLTAKYDAAKATVLDNSSTQNLNEAKTRLETAIRTAATSKQTF 134
QY 155 DNEHPNLVAYKALKTTLEORATNLEGLSSTAYNOIRNLDVLYNKASSLITKTLDPNG 214
Db 135 DQHAELVYVEELKTTLSNETATLAPYAAQYAGIKMHLGSLYDNGKAITTKTLEPVEG 194
QY 215 GTLLDSNEITTANKNINNTL--STINEQKTADALSNSFIKKVIQNEQSFVGTFTNANV 272
Db 195 DP-LTADVVMANTKIVEAIKDEVLPQENATKLADSFVKQVLYKEKITGVEAHN-KA 252
QY 273 QPSNTSFVAFSADVTPV-----NYKYARTVW-NGDEPSSRILANT-----NSI 315
Db 253 QPANTSFVGYSDITGTVTGTQTSIPNWDYAQRTIETNGDEP--RSISNTPADQTMVQPL 310
QY 316 TDVSWIYSLAGTNTKYQFSNYPSTGYLYPPYKLVKAAADANNVGLQYKLNNGVQVE 375
Db 311 SNVSWIYSLAGTCAKYTLFTYGPSTGYLYPPYKLVNTSDQMKLGEYKLN-----363
QY 376 FATSTSA-----NNTTANTPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMKNKVP 430
Db 364 -ATEPSAITFEGSEQTMNGKTPVNDINAKVTLNANKFGSNKIEFSVPA-----EKVSPM 417
QY 431 IGNIYLSNNENNAK 446
Db 418 IGNIYLSNNENNAK 433

RESULT 9

```


RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RL surface protein, pmga, of Mycoplasma gallisepticum.";
DR EMBL; L28424; AAA62416.1; -.
FW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 32.1%; Score 745.5; DB 2; Length 649;
Best Local Similarity 40.2%; Pred. No. 4.2e-27;
Matches 177; Conservative 65; Mismatches 135; Indels 63; Gaps 12;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNG-----TOLEAARMELTD 90
DB 15 IGSFVLAASCTTPTPNPPSGMGGDTNPGDGGMMNAASOEAAARMGLTT 74

QY 91 LINAKMTLASLODYAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQANTD 150
DB 75 IFDSKAKNLGLYVDYKKTONTLTAKYDAKTVLDNSSSTTQNLNEAKTRLETAIRTAATS 134

QY 151 KTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYQIRNLDVLYNKASSLITKTLD 210
DB 135 KOTFEQHAELVKVYKELKTLSTNETATLAPYADQYAGIKMHLGSLYDAGKAITKTLE 194

QY 211 PLNGTGLDLSNEITANKNINNTL--STINEQKTNADALSNSFKVKKIONNEQSFVGTFT 268
DB 195 PVEGDP-LTASAVMMANTKIVEAKIDEVLNPKENATKLAISFYKQVLVEKIKITGVEAH 253

QY 269 NANVQPSNYSFVARSADTPV-----NKKYARTVW--NGDEPSSRLANT----- 312
DB 254 N-KAQPANYSFVGSVDITGTTGTSIPNDYAQRTITNSDEP--RSISNTPADGGTM 310

QY 313 -NSTIDSVIYSLAGTNTKYQSFNSYGPSTGYLYPPYKLVKAADANNVGLYKLNNGNV 371
DB 311 AQPISNWSIYSLAGTNGAKYILEFYGPSTGYLYPPYKLVNTSDQVGLGLEKLYND--- 367

QY 372 QOVEFATSTA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMNK 426
DB 368 -----ATKPSAITFGSDQTMNGKPTVNDINVAKVTLANLFGSNKIEFSVPA-----EK 417

QY 427 VAPMIGNIYLSNENNAKDI 446
DB 418 VSPMIGNIYSSPNNNKI 437

RESULT 12
Q8RLX9 PRELIMINARY; PRT; 486 AA.
AC Q8RLX9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Hemagglutinin (Fragment).
GN VLUHA.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RC STRAIN=ts-11;
RA Markham P.F., Kanci A., Caifra G., Sundquist B., Haines P.,
RT Browning G.F.;
RL "Malp homolog in Mycoplasma gallisepticum is not essential in vitro.";
DR Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
EMBL: AY065985; AA158980.1; -.
FT NON_TER 486 486
SQ SEQUENCE 486 AA; 51844 MW; 94912DD7A09FB911 CRC64;

Query Match 31.9%; Score 740.5; DB 2; Length 486;
Best Local Similarity 37.8%; Pred. No. 5.1e-27;
Matches 174; Conservative 69; Mismatches 128; Indels 89; Gaps 12;

QY 51 PPVGSVIRLEFGCMSTTKKDNPNNG-----QTOLAEARMELTDLINAKAMTILASLOD 104
DB 52 PPSGG-----NMNGGDTNPGNGGGMDNAVQOLAAAKTALTATLLNGQTERKVLGYND 101

QY 105 YAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQANTDKTTEDNEHPNLVEA 164
DB 102 YAKIKDDLKAYIAAKEISDKSHATLQEVNNAKTRLETAIKDAANSKTSFGKKNPELKA 161

QY 165 YKALKTTLEQATNLEGLSSTAYQIRNLDVLYNKASSLITKTLDPLNGGTLDSNE-- 222
DB 162 YDALKQTTTSEMSLNQLMDANFETIKHISNLKQGRDIITATLDPITG---DGPQAM 217

QY 223 -ITANKNINVTLSINEQKTNADALSNSFKVQION-----NEQSFVGTFTNANVQPSNY 277
DB 218 VVNOTNEAIVNATSKIEDKTNATLATREFVKQTLNNANLVNE-----TNNQPPGSGY 270

QY 278 SFVAFSADV-----TPVNYKYARTVW--NGDEP-----SSRLANTNSITDVSIT 322
DB 271 SFVAYSVDLNTGVSTASNTN--NWNLAQRKVVWVGSGGRTSPFSSDANNSPALTDVSWIY 329

QY 323 SLAGTNTKYQSFNSYGPSTGYLYPPYKLVKAADANNVGLYKLNNGNVQOVEFA----- 377
DB 330 NLSGANSKYTLTTFNMGSPSTGHLVPPYKLVKSDDAQNVGLQVTLNNAKPAQRIEPAQSP 389

QY 378 -----TSTSANNTANPTPAVDEIKVAKIVLSGL 406
DB 390 SSGGTAHASDQSPRAAATETDVSASGSAQATDMSSSMNKPTPTVSDINVASVTLSDL 449

QY 407 RFGQNTIELSVPTGEGNMKVAPMIGNIYLSNENNAKDI 446
DB 450 NFGANTIIEFVPMGD---SWAPMIGNIYITSNPLVNVQI 486

RESULT 13
Q49468 PRELIMINARY; PRT; 647 AA.
AC Q49468; Q53303;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hemagglutinin homolog precursor.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93162830; PubMed=8432610;
RA Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmga, a
RT hemagglutinin of Mycoplasma gallisepticum.";
RL Infect. Immun. 61:903-909(1993).
DR EMBL; M83178; AAA02996.1; -.
KW EMBL; S55216; AAB25397.2; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match 30.8%; Score 715.5; DB 2; Length 647;
Best Local Similarity 39.2%; Pred. No. 1e-25;
Matches 174; Conservative 63; Mismatches 134; Indels 73; Gaps 13;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNG-----TOLEAARMELTD 90
DB 15 IGSFVLAASCTTPTPNPPSGMGGDTNPGDGGMMNAASOEAAARMGLTT 74

QY 91 LINAKMTLASLODYAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQANTD 150
DB 75 IFDSKAKNLGLYVDYKKTONTLTAKYDAKTVLDNSSSTTQNLNEAKTRLETAIRTAATS 134

QY 151 KTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSTAYQIRNLDVLYNKASSLITKTLD 210
DB 135 KOTFEQHAELVKVYKELKTLSTNETATLAPYADQYAGIKMHLGSLYDAGKAITKTLE 194

```
Db 135 KQFDEQHAELVKVYKELKTTLSNETATLAPYADAQYAGIKMHLISGLYDAGKAITTKTLE 194
QY 211 PLNGTGLDLSNEITFANKNINNTL--STINQKTN-----ADALNSFIKKYQIONNEOSFV 264
Db 195 PVEGDP-UTASAVMANTKIVEAIKDEVLPNPKENATKLADSLSSIVKKTIGVEE---- 249
QY 265 GTFTNANVQPSNYSFVAFSADTVP-----NYKYARRTVW-NGDEPSSRILANT--- 312
Db 250 ---AHNKAQAPANSYFVGKRYWTELLLDKQVFPNDYAGRTFTNSDEP--RSISNTPAD 304
QY 313 -----NSTDWSIYSLAGTNTKQFSPNSYGPSTGYLYFFPKLVKAADANNVGLQYKLN 367
Db 305 GQTHAQPILSNWSIYSLAGTCAKYLEFTYTGPGSTGYLYFFPKLVNTSDQVKGLEYKLN 364
QY 368 NGNVOQVEFATSTA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
Db 365 D-----ATKPSAITFGSDQTMNGKPTVNDINVAKVTLANLNFSGNKIEFSVPA--- 413
QY 423 MNKVAPMIGNIYSSNENNADKI 446
Db 414 --EKVSPMIGNMYLSSSPNNMKI 435

RESULT 14
Q9KH15
ID Q9KH15 PRELIMINARY; PRT; 656 AA.
AC Q9KH15;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Adhesin pMGAL.2.
GN pMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pMGA multigene family of Mycoplasma
gallisepticum strain HS."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF275312; RAF91413.1; --
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E073FC617E1 CRC64;

Query Match 30.3%; Score 704; DB 2; Length 656;
Best Local Similarity 41.0%; Pred. No. 3.5e-25;
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

QY 71 DANPNNGOT-----QLEARMELTDLINAKAMTSLASQDYAKTEASLSAYSAETVN 123
Db 63 DINPGGQNMDSAAQELTAARTALTSLSKKNANVEMYSDYAKIQNTLIAAYTTAEQTS 122
QY 124 NNLNATLEOLKMAKNLLESAINQANTKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 183
Db 123 QNSATLQOVKNATGALQTAINTANSKQKFDHSDNLLMSYKNLMLATLAKKATVMTLK 182
QY 184 STAYNQIRNRLVDLYNKASSLTTKTLDPLNGTGLDLSNEITFANKNINNTLS--TINEQK 241
Db 183 DPKYSAILDQINGVSSKGEELVQHTLDPSV-CIVPAANTITEITKIEBIVISEKTLQOK 241
QY 242 TNADALS--SFI--KKVIONNEQSFVGTFTNANVQPSNYSFVAFSADTVP----- 289
Db 242 NNADQAFYQSFTLDKTKLENVEDA-----KKMQQPANSYFVGSDVTGTSQGETRIP 295
QY 290 NYKYARRTVWNGDEPSSRILANTNS-----ITDWSIYSLAGTNTKQFSPNSYGPS 341
Db 296 MNWFAQALFTSGNOPTKATTTGCDGOSTAKPLSDVSWIYSLAGTCAKYLEFTIYGPS 355
QY 342 TGYLYFFPKLVKAADANNVGLQYKLN-NGNVOQVEFATSTANNTANPTPAVDEIKVAK 400
Db 356 TGMWLYFFPKLVKAND--DVGLQYKLSNETLTPILFEGCT-----TTNGPAATVENINAK 409
```

```
QY 401 IVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYVLSNENNADK 445
Db 410 VRLTGLAFGKNTIEFSVP-----MSKVAPMIGNMYITSSDTEYNK 449

RESULT 15
O05122
ID O05122 PRELIMINARY; PRT; 703 AA.
AC O05122;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Haemagglutinin.
GN pMGAL.9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
surface protein, pMGA, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90714; AAB50154.1; --
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CBC5 CRC64;

Query Match 30.1%; Score 700.5; DB 2; Length 703;
Best Local Similarity 34.8%; Pred. No. 5.5e-25;
Matches 170; Conservative 77; Mismatches 126; Indels 115; Gaps 12;

QY 53 VGSTVIRLEFGCMSTIKKDANP-----NNGOT-----OLEAA 84
Db 15 IGSFVMLAAASCTSATIPTLNPPEPKPDPMPNPPSGMGNGNTNPGMDTAAQELASA 74
QY 85 RMELTDLINAKAMTSLASQDYAKIEASLSAYSAETVNNNLNATLEOLKMAKNLESAL 144
Db 75 KAALTTLNRESEKVGLYVDYAKIKADLTSAVTAKTSDSTSLVQVKTATSKLQTAI 134
QY 145 NQANTDKTTFDNEHPNLVEAYKALKTTLEOR-ATNLEGLSSTAYNOIRNRLVDLYNKASS 203
Db 135 DKRAADKQKFFQDHDLLMPYSELATYLSQKNATVL--LNQPKYSAILNKINSIYAQGE 192
QY 204 LITKTLDPLNGTGLDLSNEITFANKNINNTLS--TINEQKTNADALSFIKKYQIONNEQ 261
Db 193 VVIRTLDPVS-GAIPTAASITKVNDINKAISENQLKPKKDADAFANYQFFKL---DKT 248
QY 262 SFVGFTNANVQPSNYSFVAFSADTVP-----NYKYARRTVWNGDEPSSRILANTN 313
Db 249 KIMGSTNMKQPNYSFVGYSVGTGMSGOTTIPNNWFAQRIWSSGAPRAPLASQTE 308
QY 314 S----- 314
Db 309 TPQAEPTPMSPAPQVEPQAQOGSDSPKQASQTEVSPTPAAEVQAQADTEQPATSQGTP 368
QY 315 ITDWSIYSLAGTNTKQFSPNSYGPSTGYLYFFPKLVKAADANNVGLQYKLNNGVQV 374
Db 369 LTDVSWIYSLGTDVKYTFYTFNYFGPSMAYLYFFPKLVKSD--SVGLQYKLNNNNPAL 426
QY 375 EPATSTANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNI 434
Db 427 NFGSETNAN-----GPAASVDNINVAKVNLNANFNCGENTIEFSVP-----MNKVAPMIGNM 477
QY 435 YLSSNENN 442
Db 478 YITSDVAN 485
```

Fri Oct 10 11:43:09 2003

us-09-147-052-2.rspt

Page 8

Search completed: October 8, 2003, 17:06:59
Job time : 43.2516 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:35 ; Search time 8.2108 Seconds
(without alignments)
2611.703 Million cell updates/sec

Title: US-09-147-052-2

Perfect score: 2324

Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNENNADKIPGRRPCTFL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	314	13.5	865	1 VGLB_HSVMD	P18538 marek's dis
2	154.5	6.6	682	1 NISP_LACIA	Q07596 lactococcus
3	152	6.5	2660	1 YEEJ_ECO57	Q848v7 escherichia
4	146.5	6.3	857	1 AR56_CANAL	P78586 c arg5,6 pr
5	146.5	6.3	1085	1 CUY7_SCHPO	P24339 schizosacch
6	146	6.3	719	1 YM41_YEAST	Q03213 saccharomyc
7	143	6.2	3178	1 YS89_CAEEL	Q09624 caenorhabdi
8	141	6.1	1314	1 SW11_YEAST	P09547 saccharomyc
9	140	6.0	1381	1 YBE7_YEAST	P34216 saccharomyc
10	138	5.9	444	1 PST1_YEAST	Q12355 saccharomyc
11	138	5.9	995	1 YI09_YEAST	P40442 saccharomyc
12	138	5.9	1010	1 YKK1_CAEEL	P34278 caenorhabdi
13	138	5.9	1260	1 ALS1_CANAL	P46590 candida alb
14	137	5.9	1140	1 YM96_YEAST	Q04893 saccharomyc
15	136.5	5.9	956	1 YEF3_YEAST	P32618 saccharomyc
16	135	5.8	1271	1 Y338_MYCGE	P47580 mycoplasma
17	133.5	5.7	903	1 VGLB_HSV1F	P06436 herpes simp
18	133.5	5.7	904	1 VGLB_HSV1F	P10211 herpes simp
19	133.5	5.7	904	1 VGLB_HSV1P	P08665 herpes simp
20	133.5	5.7	1018	1 FNBA_STRAU	P14738 staphylococ
21	132.5	5.7	457	1 MESE_LEUME	Q10419 leuconostoc
22	132.5	5.7	1672	1 PMPB_CHLMU	Q9PJY2 chlamydia m
23	132	5.7	1902	1 P3P_LACLC	P15292 lactococcus
24	131.5	5.7	1164	1 BAG_STRAG	P27951 streptococc
25	131	5.6	837	1 ROD1_YEAST	Q02805 saccharomyc
26	131	5.6	1433	1 CAT8_YEAST	P39113 saccharomyc
27	130.5	5.6	704	1 MSN2_YEAST	P33748 saccharomyc
28	130.5	5.6	800	1 INLA_LISMO	P21416 listeria mo
29	130.5	5.6	1159	1 N124_SCHPO	Q09904 schizosacch
30	130.5	5.6	1161	1 DAN4_YEAST	P47179 saccharomyc
31	130.5	5.6	1902	1 P2P_LACLC	P15293 lactococcus
32	130	5.6	425	1 YBY0_YEAST	P38272 saccharomyc
33	129	5.6	1251	1 RBP2_PLAYB	Q00799 plasmodium

34 129 5.6 1790 1 USOL_YEAST P25386 saccharomyc
35 128.5 5.5 750 1 YKS7_YEAST P34231 saccharomyc
36 128.5 5.5 1858 1 P3K2_DICDI P54674 dictyosteli
37 128.5 5.5 2334 1 WAPA_BACSU Q07833 bacillus su
38 128.5 5.5 3712 1 LNA_DROME Q00174 drosophila
39 128 5.5 1117 1 YN96_YEAST P53753 saccharomyc
40 127.5 5.5 544 1 FLKG_BUCAP Q8K9K0 buchnera ap
41 127 5.5 762 1 SLAP_ACEKI P22258 acetogenium
42 127 5.5 979 1 P115_MYCHR P41508 mycoplasma
43 127 5.5 1142 1 GIN4_YEAST Q12263 saccharomyc
44 127 5.5 1643 1 OMPB_RICPR Q53020 r outer mem
45 126.5 5.4 904 1 VGLB_HSVIK P06437 herpes simp

ALIGNMENTS

RESULT 1
VGLB_HSVMD
ID VGLB_HSVMD STANDARD; PRT; 865 AA.
AC P18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_FaxID:33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293086; PubMed=2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Binns M.M., Doel T., Milne B.;
RT "Nucleotide sequence and characterization of the Marek's disease
virus homologue of glycoprotein B of herpes simplex virus.";
RL J. Gen. Virol. 70:1789-1804(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL: D13713; BAA02866.1; -
InterPro: IPR000234; Glycoprot_B.
Pfam: PF00606; Glycoprotein_B; 1.
ProDom: PD000693; Glycoprot_B; 1.
Signal: Glycoprotein; Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 865 GLYCOPROTEIN B.
FT DOMAIN 22 682 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 683 700 POTENTIAL.
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 732 752 POTENTIAL.
FT DOMAIN 753 865 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 98091 MW; B30E93C1AC65C6C3 CRC64;
Query Match 13.5%; Score 314; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVLYGTNSSPSTQNTVTSREVVSSVOLSEESTFYLCPPVGSTVRL 60
|||||

Db 1 MHYFRRCIPFLVILYGTNSSPSTQNTSVTSREVSVQLSEESTFYLCPPPVGSIIVRL 60
QY 61 E 61
Db 61 E 61

RESULT 2
NISP_LACLA STANDARD; PRT; 682 AA.
AC Q07596;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nisin leader peptide processing serine protease nisp precursor
DE (EC 3.4.21.-).
GN NISP.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIZO R5;
RX MEDLINE=93239683; PubMed=8478324;
RA van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,
RT Kuipers O.P., de Vos W.M.;
RT "Characterization of the Lactococcus lactis nisin A operon genes
RT nisp, encoding a subtilisin-like serine protease involved in
RT precursor processing, and nlsr, encoding a regulatory protein
RT involved in nisin biosynthesis.";
RT J. Bacteriol. 175:2578-2588(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6F3;
RX MEDLINE=94213458; PubMed=8161176;
RA Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,
RA Hammelmann M., Entian K.-D.;
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
RT 6F3.";
RL Appl. Environ. Microbiol. 60:814-825(1994).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=95357326; PubMed=7630881;
RA Siezen R.J., Rollem H.S., Kuipers O.P., de Vos W.M.;
RT "Homology modelling of the Lactococcus lactis leader peptidase Nisp
RT and its interaction with the precursor of the lantibiotic nisin.";
RL Protein Eng. 8:117-125(1995).
CC -!- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC -!- PATHWAY: Nisin biosynthesis; last step.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L11061; AAA25200.1; -
CC DR EMBL; X76884; CAA54210.1; -
CC DR PIR; S44131; S44131.
CC DR HSP; P29600; IGCI.
CC DR MEROPS; S08.059; -
CC DR InterPro; IPR006192; LPXTG.
CC DR InterPro; IPR00209; Peptidase_S8.
CC DR Pfam; PF00082; Peptidase_S8; 1.
CC DR PRINTS; PR00723; SUBTILISIN
CC DR PROSITE; PS00136; SUBTILASE ASP; 1.
CC DR PROSITE; PS00137; SUBTILASE HIS; 1.
CC DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.

DR PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
KW Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 195
FT CHAIN 196 655
FT NISIN LEADER PEPTIDE PROCESSING.SERINE
FT PROPEP 656 682
FT ACT_SITE 259 259
FT ACT_SITE 306 306
FT ACT_SITE 512 512
FT SITE 652 656
FT MOD_RES 655 655
FT CONFLICT 500 500
SQ SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;
Query Match 6.6%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.22;
Matches 75; Conservative 61; Mismatches 130; Indels 91; Gaps 13;
QY 120 ETVNNLNATLEQLKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLEQATNL 179
Db ELINHSNALLSTEGSTTDSINLGAQSPAVKSTTRTE----LDVTGAATLLQTSVQK 93
QY 180 EGLSSYAYQIRNNLVLDLYNKASLLTKTLDPLNGTGLDLSNEITANKNNITLSTINE 239
Db EMKVSQETQVSSE----FSKRDVNTKEAVPVSKDELLEQSEVVVSTSIQKN-KILDN 148
QY 240 QKTNADALNSFLKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKARRTV- 298
Db KKRAFNVTSSPLIKEKPSNKSADSGVIDNS-----ASPLSYRKKEVVS 193
QY 299 -----WNGDEPSSRLANTNSITDWSIYSLAGTNTKYQFSNYGPS 341
Db LROPLANKQVEAQPILLISNSSEKSKASYTNSHDFWYQW-----DMKY---VTNNGES 243
QY 342 TGYLYPEYKLVKRAADANNYQVLYKLNNGVQVQVEFATPSAN-----NT 385
Db YA-LYOPSKKI-----SVGI-----IDSGIMEHPDLSNLGNYFKNLVPGKGFDEE 293
QY 386 TAMPPTAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKMKVAPMTIG-NIYLSNEN 441
Db TGNPSDIVD-----KMGHTEVAGQITANGNLLGVAPGIVTVIYRVEGEN 338
RESULT 3
YEEJ_ECO57 STANDARD; PRT; 2660 AA.
AC Q8X8V7; Q8X2B9; Q8X2C0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yeeJ.
GN 23135 OR ECS2775/ECS2776.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT *Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -|- SIMILARITY: Contains 16 Big-1 domains.
CC -|- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.
CC -|- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 1315.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005423; AAG57041.1; -
DR EMBL; AP002559; BAB36198.1; ALT_FRAME.
DR EMBL; AP002559; BAB36199.1; ALT_FRAME.
DR InterPro; IPR003344; Big_1.
DR InterPro; IPR003335; Intimin.
DR Pfam; PF02369; Big_1; 16.
DR PRINTS; PR01369; INTIMIN.
DR SMART; SM00634; BID_1; 16.
DR SMART; SM00089; PKD; 8.
KW Hypothetical protein; Repeat; Complete proteome.
FT DOMAIN 738 834 BIG-1 1.
FT DOMAIN 840 929 BIG-1 2.
FT DOMAIN 931 1033 BIG-1 3.
FT DOMAIN 1042 1132 BIG-1 4.
FT DOMAIN 1134 1236 BIG-1 5.
FT DOMAIN 1245 1335 BIG-1 6.
FT DOMAIN 1337 1439 BIG-1 7.
FT DOMAIN 1448 1539 BIG-1 8.
FT DOMAIN 1548 1652 BIG-1 9.
FT DOMAIN 1653 1750 BIG-1 10.
FT DOMAIN 1751 1855 BIG-1 11.
FT DOMAIN 1856 1957 BIG-1 12.
FT DOMAIN 1963 2056 BIG-1 13.
FT DOMAIN 2065 2156 BIG-1 14.
FT DOMAIN 2157 2252 BIG-1 15.
FT DOMAIN 2254 2355 BIG-1 16.
SQ SEQUENCE 2660 AA; 280062 MW; 01EB92A08F5C09D2 CRC64;

Query Match
Best Local Similarity 6.5%; Score 152; DB 1; Length 2660;
Matches 108; Conservative 70; Mismatches 188; Indels 104; Gaps 20;

QY 12 LIVILYGNSSSTONVTSREVS-----SVQLS-----EESTFYLCPPPPGVGVIRLRCG 64
DB 1131 IVNIAPDASNAQVTLNIPAAQVVTNNSDSVQLTATVKDPSNH-----PVAGITV----- 1179
QY 65 MSITKDPANPNNGQQLAARAMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVYNN 124
DB 1180 -----NFTMPQDVAANFTLEN--NGIAITQANGEAHVTLKGGKAGHTVTTVLGN 1227
QY 125 NIATLEQLKMAKTWLESAINQANT-----DKTT-----FDNEHPNLVEAYK 166
DB 1228 N--NASDAQVTEVADKDSAVVVLQTSKAEIIGNGVDETTLTATVKDPFDN-----AVK 1279
QY 167 AKTTLEQATNLEGLSSTAYNIQNNLVLYNKASSLITKLDPLNGGTLT--DSNETT 224
DB 1280 DLQVTF--STN--PADTQLSQSKSNTND-----SGVAEYTFK-----GTVLGVHTAEAT 1324
QY 225 TANKNNINLTSLINEQKINADALSNTFKKVIQNNEQSFVGTFTNANVPQSNYSFVAFSA 284
DB 1325 LPNGNNDKIVNIAPDASNAQVTLNIPAAQVVTNNSDSVQLTATVKD--PSNHPVAGITV 1382

```

```

QY 285 DVTVPVNYKYARTVWNGDEPSRIILANTNSITDVSIIYLACTNTKYQFSFNSYGPSTGY 344
DB 1383 NFTMPQDVAANFTL-----ENNGIAITQANGEAHVTLKGGKAGHTT-VTATLSNNNTSDSQ 1437
QY 345 LYFPYKLVKAADANNVGLQYKLNNNGVQOQVEPAT-----STSANN 384
DB 1438 ---PVTFFVADKTSALVLIQISKNETGNGVDVSATLTAIVKQDFDNEVNNLPVTFSTASSG 1494
QY 385 TTANPTPA-VDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGN 433
DB 1495 LTLTPGESNTNESGIAQATLAGVAFGEQTVTASLANNGASDNKTKVHFIGD 1544

RESULT 4
AR56_CANAL
ID AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ARG5,6 protein, mitochondrial precursor (Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5,6.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -|- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
CC -|- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -|- PATHWAY: Arginine biosynthesis; second step.
CC -|- PATHWAY: Arginine biosynthesis; third step.
CC -|- SUBCELLULAR LOCATION: Mitochondrial.
CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -|- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X98880; CAAG7383.1; -
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; Acglutkinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinas; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdehyde_dh; 1.
DR Pfam; PF02774; Semialdehyde_dhc; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARG; 1.
DR Oxioreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
KW

```

KW Mitochondrion; Multifunctional enzyme; Transit peptide.
 FT TRANSIT 1 ? ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? ? ACETYLGUTAMATE KINASE.
 FT CHAIN ? 857 N-ACETYLGUTAMATE KINASE.
 FT REDUCTASE.
 FT ACT_SITE 669 669 BY SIMILARITY.
 FT SEQUENCE 857 AA; 95066 MW; AE2CEADBF8C4C71 CRC64;
 Query Match 6.3%; Score 146.5; DB 1; Length 857;
 Best Local Similarity 22.5%; Pred. No. 0.78;
 Matches 122; Conservative 71; Mismatches 186; Indels 163; Gaps 32;
 QY 7 NCIFFLI-----VILYGNSSPTQNTVSREVSSVQSEESTYLCPPPGVSTVIR- 59
 Db SCIAFLYHVLGVIPVILHGT--GPOINELLENEGEPEYIDGIRIT-----NPKTMEVVR 156
 QY 60 --LEFGCMSTTKD-----ANPNNGOTQLEAARWELTDLINAKAMTLASIQDYAK--IEA 110
 Db 157 CFLEONRLVTALEKIGVHARP-----ITAGVFEAEYLDKDKYQVLGVKITSYVKNKSPVEA 210
 QY 111 SLSSA-----SEATVNN---NLNATL-----EQLKMAKTNLESAINOANTDK-- 151
 Db 211 AINSGLYPLITSLAETSSGOLLNVNADVAAGELAREPEPLKIVYLNEKGGIINGNTGEKV 270
 QY 152 --TFDNEHPNLV-EAYKALKTTLEOR-----ATNLEGLSSTAYNOIRNNLVLYNKASS 203
 Db 271 SAINLDEEYEDLKESWVKYGTALKIKREIHDLLQHLPRSSVA-----IIDVDLQKE 323
 QY 204 LITKTLDPNGTGLDLSNEITTANKNIN--NTLSTINEQKTNADALSNTFKKVIQNEQS 262
 Db 324 LFTDS-----GAGTL-----IRRGYRLINRSLRDFG-----NPDLNALLR----- 361
 QY 263 FVCTFTNANVQPSNYSFVAFAADVTPVNYKYARTVWNGDEPSSRILANT----- 312
 Db 362 -----DPEIKTKGVSVASYKLFLDSVQFKS-----YGDPE-LEVLAIVEQNDKIPKL 408
 QY 313 -----NSITD-----VSWIYSLAGTNTKQFSPFSNFG-PSTGYLYPP 348
 Db 409 DEFLLSKTGLWNLVNTVINAKKQYQOLCVVNVENDANLPMWFSKDSGFAKNGQILFW 468
 QY 349 Y-----KLKKAADANNVG--LOYKLNG-----NYQVVE-FATSTSANNTANPTPAVD 394
 Db 469 YGLNIDEASKLKEFDSSIGSSLSKESGVFTSAQKRGPHHSIVRNT--NPNPPLS 526
 QY 395 EKVA---KIVLSGLR--FCONTIELSVPTGEGNMKNKVPMTGNIYLSNENNADKIPGY 449
 Db 527 EGQTERKKVALTGARGYTGQNLIKLI-----DNHPYLDISYVSSRELEGQKLOGY 577
 QY 450 RR 451
 Db 578 NK 579
 RESULT 5
 CUT7_SCHPO
 ID CUT7_SCHPO STANDARD; PRT; 1085 AA.
 AC P24339.
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein cut7.
 GN CUT7 OR SPAC25G10.07C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91015362; PubMed=2145514;
 RA Hagan I., Yanagida M.;
 RT "Novel potential mitotic motor protein encoded by the fission yeast
 cut7+ gene.";

RL Nature 347:563-566 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sources J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakert G., Aert R., Robben J., Grymonprez B.,
 RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 CC Nature 415:871-880 (2002).
 CC -1- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
 CC FORM A SHORT SPINDLE THAT ELONGATES TOWARD THE NUCLEUS AT
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. B1MC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X57513; CAA40738.1; -;
 CC EMBL: Z70691; CAA94636.1; -;
 CC FIR: T38378; T38378.
 CC HSSP: P17119; 3KAR.
 CC GeneDB: SPombe; SPAC25G10.07c; -;
 CC InterPro: IPR001752; kinesin_motor.
 CC Pfam: PF00225; kinesin; 1.
 CC PRINTS: PR00380; KINESINHEAVY.
 CC SMART: SM00129; KISC; 1.
 CC PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 CC PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 CC Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).
 FT DOMAIN 897 955 COILED COIL (POTENTIAL).
 FT NP_BIND 159 166 ATP (BY SIMILARITY).
 FT REPEAT 987 998
 FT REPEAT 999 1010
 FT MOD_RES 1011 1011
 FT PHOSPHORYLATION (BY CDC2) (BY
 FT SIMILARITY).
 FT SASNPKRREPTIDTGYDGRSDTNSPT ->
 FT LRAIGNDYSLLLLL (IN REF. 1).
 SQ SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;

Query Match 6.3%; Score 146.5; DB 1; Length 1085;
 Best Local Similarity 20.2%; Pred. No. 1.1;
 Matches 110; Conservative 89; Mismatches 204; Indels 141; Gaps 25;

QY 17 YGTNPSPTQNTV-----GREVVSV-----QLSEESTFYLCPPVPGSTVIR 59
 DB 554 YETNEAKITVATDLSQYRESKEYIASLYEKLDRTERNKENNFW-----N 602

QY 60 LEFGCMSTYK-----KDANPNNGO--TOLEARMELTDLINAKA--MTLASL-----QDYAKIE 109
 DB 603 LKFNLTMLRSPHSGSTDTNGYFTLLDNFNASMEELLNTHSNQLLISMTKITEHFQSLD 662

QY 110 ASLSSAYSEATVNNLNATLQKMAKTNLESAINQANTDKTTFDNEHPN-----LVZA 164
 DB 663 EALQSARSCAVPNSLDLIVSELKDKNSLLDALEHSLQDISMSQKLGNGISSLEL 722

QY 165 YKALKT-----LEQRATNLEGLSSTAYNQIRNNLVLDLYNKASSLITKTLPLNGTLLDSN 221
 DB 723 QKDMKESYRQLVQLRSLYNLQHTHEESQKELMYGVRNDIDALVKTCTTSLNADADILSD 782

QY 222 EIT-----TANKNINLTSTINEQ--KTNADAL-----SNSFIKK---VI 256
 DB 783 YISDQSKFESKQOQDLIANIGIVSNFLQEQNESLYTKADILHSHLNDTNSNIRKANEIM 842

QY 257 QNNEQSFVCTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNG-----DEPSSRLANT 312
 DB 843 NNRSEEL-----RNA-----ASQAEIVGANKERIQKTENGSQLDSSKAIHSN 889

QY 313 NSITDVSMTYSLA-----GTNTKYQ-----FSFSGYGPSTGYLYFPYKLVKA 354
 DB 890 RSMYD-----HCLALAESQKQGVNLEVQTLDRLLQKVEHSEDNTEKHOQL---IDLLES 942

QY 355 ADANNVGL--QYKLNNGVQOVE-----FATSSANNIT-----ANPTPAVDEIKV 398
 DB 943 LVGNNDNLDSIKTPHTLQKTDHVLKGTISLANHTNELLGLGDESLCNLETTIEDTSL 1002

QY 399 AKTVLSGLREGQNTI-----ELSVPTGEGNMKNVAPMIGNIYLSNENNADKIPG 448
 DB 1003 VLETTGDPSPKRELPAFSPWTRDSSLIKETTLNLDSDKKFVRETYTSSNOTNEPDV-- 1060

QY 449 YRRP 452
 DB 1061 YDKP 1064

RESULT 6
 YM41_YEAST STANDARD; PRT; 719 AA.
 ID YM41_YEAST STANDARD; PRT; 719 AA.
 AC Q03213;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 79.4 kDa protein in Ald2-DDR48 intergenic region.
 GN YMR172W OR YMR010.02.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 XII.";
 RL Nature 387:90-93(1997).
 CC -1- SIMILARITY: LOW, TO YEAST MSN1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL; Z49808; CAA89905.1; --
 DR PIR; S53119; S53119.
 DR TRANSFAC; T04601; --
 DR SGD; S0004783; HOT1.
 DR GO; GO:0000228; C:nuclear chromosome; IDA.
 DR GO; GO:0006972; P:hyperosmotic response; IGI.
 KW Hypothetical protein.
 SQ SEQUENCE 719 AA; 79415 MW; 4652BE93743D5A54 CRC64;

Query Match 6.3%; Score 146; DB 1; Length 719;
 Best Local Similarity 21.4%; Pred. No. 0.67;
 Matches 101; Conservative 69; Mismatches 174; Indels 128; Gaps 20;

QY 19 TNSPSTQNTVREVSVSVOLSEESTFYLCPPVPGSTVIRLEFGCMSTKDKDANPNNGQ 78
 DB 150 TNDSPS--NEISTDQLKIFQMDMS-----ARMIEMESFNKLSNKIAEQNTMV 197

QY 79 TOLEARMELTDLINAKAMTLA-----SLDYAKIE-----ASLSAYSEAEVNN 124
 DB 198 LNLKODNYKVMNKLNTLLKLVQAPSARPSTNNAQNKLAIELLSISAVSSAYLQRMQNG 257

QY 125 NLNATLEQLKMAKTNLESAINQ-----ANTDFTDNEHPNLVEAYKALKTTLEOR 175
 DB 258 SGRQHTADLCTGDSNTHSGINQHRITNGTIDVNTAQLNNQFSN-----ALNTILPDQ 311

QY 176 ATNLEGLSTAYNQIRNNLVLDLYNKASSLITKTLPLNGTLLDSNEITTANKNINLT- 234
 DB 312 QHN-----RNNVSONIQ--SLPNROLGPVN-----TOANQNSQVLI 348

QY 235 -STINEQKTNADALS--NSFIKKVIONNEQSFV-----GFTNANVQPSNYFVA---FSA 284
 DB 349 HNTNTHQQVNRSPISFPNASTDKPFLNPGIKRRRNTQSNNNASTNDHASAAQKPSA 408

QY 285 DVTVPVNYKARTVWNGDEFS--SRILANTNSTIDVSIYSLAGTNTKYQFSFSPNGPST 342
 DB 409 -LSPLTNSHNTSMNTNTHSSIHSGVTSASNSFHDLN-----SLNNGFTT 453

QY 343 GYLIFYPKLVKADANNVGLQYKLNNGVQOVEFATSTANNTANTPAVDEIKVAKIV 402
 DB 454 ALSPLSLALDNASFPNPQNVIPPIINNTQQPLSFLINQDSTTS----- 498

QY 403 LSLREGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKIPYRRPCT 454
 DB 499 -----EL-LPSGKSGVNT-----NIV---NNRRASTLPSYKPKMT 529

RESULT 7
 YS89_CAEEL STANDARD; PRT; 3178 AA.
 AC Q09624; Q09625; Q969D4;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein ZK945.9 in chromosome II.
 GN ZK945.9/ZK945.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilkison-Sproat J.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Durbin R.;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC	EMBL; X12493; CAA31013.1; -	
DR	EMBL; U33335; AAB68089.1; -	
DR	PIR; S05728; TNYR6.	
DR	PDB; 1KKX; 04-DEC-02.	
DR	PDB; 1KN5; 04-DEC-02.	
DR	TRANSFAC; T01279; -	
DR	SGD; S0005937; SW1.	
DR	InterPro; IPR001606; ARID.	
DR	Pfam; PF01388; ARID; 1.	
DR	SMART; SM00501; BRIGHT; 1.	
KW	Transcription regulation; Activator; DNA-binding; Nuclear protein;	
KW	zinc-finger; 3D-structure.	
KW	DOMAIN 5 65	ASN/THR-RICH.
FT	DOMAIN 337 385	GLN-RICH.
FT	2N_FING 1241 1258	C4-TYPE.
FT	SEQUENCE 1314 AA; 147938 MW; F442D5A82013CD8D	CRC64;
SQ		

Query Match 6.1%; Score 141; DB 1; Length 1314;
Best Local Similarity 19.9%; Pred. No. 2.6;
Matches 79; Conservative 63; Mismatches 154; Indels 100; Gaps 17;

```

20  NSPSPQNTVTSREVWSSVOLSEBESTYFLCPPVPVGSTVIRLRFPGCMSITKKDANPNNGOT 79
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
13  NNTTTTTTTTTNTNTNNNTNNNN-----PANNTNNNNSTGHSSTNNNTNNNTNT 65
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
80  QLEAARMEITDLINAKAMTLASLODYAKTEASLSAYSEATVNNNLNATLEQLKMAKTN 139
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
66  GASG-----VDDFQN-----FFDKPFPQNDLSDNNNSNNNNNNNN--NTVASSTN 111
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
140 LES---AINQANTDKTT-----FDNEHPNLVEAYKALKTTLEQRATNIEGLSSTAYNQ 189
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
112 FTSPTAVVNAAPANVTGKKAANFIQSPQSPNPDYSSNSN-----TNLSLSPQAI-L 165
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
190 IRNNLDLYNKASLLTKTLDPLNGGTLLDSNEITTANKNNNTLSTINPQKTNADALSN 249
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
166 AKNSIID-----SSNLPLQAQQQLYGG-----NNNNNSTGIANDVITPHFITN 209
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
250 SFIKKVIQNEQSFVCTFTNANQVP--SNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRI 308
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
210 --VQISQNSSSSTPN--TNSNSTPNANOQLFPNNSASN-----NGLNLSNQL 254
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
309 LANTNSITDVSYITSLAGNTKQFSFSNPGPSTGYLYFPYKLVKAADANNVGLQYL-- 366
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
255 ISN-----YAASNSMDRSSASNEVPNT-----SDNNNSNNNNHNRN 292
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
367 -----NNGNVQOOFATSTSANNTTANPTPAVDE 395
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
293 NSNNKTSNNNNVTAVPAATPANTNNSTSNANTVFSE 328
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

RESULT 9

YBE7_YEAST	STANDARD;	PRT; 1381 AA.
ID	YBE7_YEAST	
AC	F34216;	
DT	01-FEB-1994 (Rel. 28, Created)	
DT	01-OCT-1994 (Rel. 30, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Hypothetical 130.8 kDa protein in SEC17-QCR1 intergenic region.	
DE	YBL047C OR YBL0520 OR YBL0501.	
GN	OS Saccharomyces cerevisiae (Baker's yeast).	
OS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.	
OX	NCBI_TaxID=4932;	
RN	[1]	
RP	SEQUENCE OF 1-961 FROM N.A.	
RC	STRAIN=5288C;	

```

RX MEDLINE=95176707; PubMed=7871888;
RA de Wergifosse P., Jacques B., Jonniaux J.-L., Purnelle B., Skala J.,
RA Goffeau A.;
RA "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
RT chromosome II reveals homologues to bacterial proline synthetase and
RT murine alpha-adaptin, as well as a new permease and a DNA-binding
RT protein.";
RT RT
RT Yeast 10:1489-1496(1994).
RN [2]
RN SEQUENCE OF 579-1381 FROM N.A.
RC STRAIN=S288c;
RC MEDLINE=94205256; PubMed=8154187;
RA Scherens B., el Bakkoury M., Vierdeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIP1 and SEC17 genes.";
RT RT
RT Yeast 9:1355-1371(1993).
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC CC
CC -1- SIMILARITY: Contains 3 EH domains.
CC -1- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z35808; CAA84867.1; -
DR EMBL; X78214; CAA55048.1; -
DR EMBL; Z23261; CAA80797.1; -
DR PIR; S45781; S45781.
DR SGD; S0000143; EDEL1.
DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomyces); IDA.
DR GO; GO:0005935; C:bud neck; IDA.
DR GO; GO:0005934; C:bud tip; IDA.
DR GO; GO:0006897; P:endocytosis; IMP.
DR InterPro; IPR000261; EPS15_homology.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00036; efhand; 2.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00054; EFH; 2.
DR SMART; SM00027; EH; 3.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS50031; EH; 3.
DR PROSITE; PS50030; UBA; 1.
KW Hypothetical protein; Repeat.
FT DOMAIN 14 113 EH 1.
FT DOMAIN 135 227 EH 2.
FT DOMAIN 277 366 EH 3.
FT DOMAIN 1338 1380 UBA.
SQ SEQUENCE 1381 AA; 150783 MW; 626FD261DCBA7D99 CRC64;

```

Query Match	6.08;	Score 140;	DB 1;	Length 1381;
Best Local Similarity	19.5%;	Fred. No. 3.2;		
Matches 119;	Conservative	56;	Mismatches 212;	Indels 224;
				Gaps 22;

Qy	19	TNSPSTQNTSREVYSSVOLSEESTFYLCPPP-----	52
		: :	
Db	487	TPTSNEFQSITKEEPEEQLEORESSDTFSQAQPPVKPHASSPVKRTASTTLPOVPNFVSF	546
		: :	
Qy	53	-----VGSTVIRLEFGCMSITTKDANPNQC-----	88
		: :	
Db	547	SWPAGAATSAATGAAGVAAALGASAFSRSSNNAFKNQDLFADGEASAQUSNATTEM	606
		: :	
Qy	89	TDLINAKMTLASLQDYAKT-EASLSAYSSEAFVN-----	137
		: :	
Db	607	ANLSN----QVNSLSKQASITNDKKSRAQELKRVTEMKNSIQIKLNNLRSTHDQNVKQT	662
		: :	
Qy	138	TNLESINQANTDKTTFDNEHPNLVBAYKALKTTTLEQRATNLEGLSSTAYNQRNNLVLDL	197
		: :	
Db	663	EOLEAOLVQVKNKNETLAOLAYSEANYIAAESKLNELTDLQOE-SOTKNAELKEQOITNL	731
		: :	

```

QY 198 YNKASLTKTKDPLN-----GGTLLDSNEITTAN----- 227
D  : : : : :
Db 722 NSMTASLQSQ----LNEKQQQVKQERSMVDVNSKQLELNQVTVANLQKIDGLGEKISVY 777
QY 228 ----KNINNTLSTINEQKTADA----LSNSFI-----KVIQNEQ-----SFV 264
D  : : : : :
Db 778 LTKQELNDYQKTVEEQAQQAQYQDLSNKDQTLTDREKLEERNQIEQENLYHQHV 837
QY 265 GTFTNA--NVQPSNYSFVAFSADVTVPVNYKYARITVWNGDEPSSR----- 307
D  : : : : :
Db 838 SKLQEMFDDLSQRKASFKAQDELKERIEVAN----NVRLSERQMLANGQLPEDAKD 893
QY 308 ILANTNSITDYSWISYLAGTKYQFSNYPSTGPGSTGYLYFYKLYKAA----- 355
D  : : : : :
Db 894 IIAKSASNTD-----TTTKEATSRGNVHEDT-----VSKFVETTVENSNLNVNRVK 939
QY 356 -----DANNVGLQYKLNNGVQVQEFATPSTSANNTTANP--TPAYDEIK 397
D  : : : : :
Db 940 DEEKTEKTESDVFDRDPTGTSQSDSENANTN-----NGTQSGNETANPNLTETLSD-- 992
QY 398 VAKIVSLGRFQNTIELSVPTGEGNMKNVA-----PMIGNIYLS 437
D  : : : : :
Db 993 -----RFGDLNEXGIPRSQSLTSSVANNAPQSVRDDVPELTPETLEERTINNTANR 1043
QY 438 SHENNADKIPG 448
D  : : : : :
Db 1044 DNTGNLSHIPG 1054

RESULT 10
PST1_YEAST STANDARD; PRT; 444 AA.
ID AC Q12355;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protoplast secreted protein 1 precursor.
GN PST1 OR YD055W OR D4214 OR YD9609.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96381250; Pubmed=8789263;
RA Brandt P., Ramlow S., Otto B., Bloeker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
of Saccharomyces cerevisiae chromosome IV."
RL Yeast 12:85-90(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION.
RX MEDLINE=99251092; Pubmed=10234784;
RA Pardo M., Montecoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
RT "Two-dimensional analysis of proteins secreted by Saccharomyces
cerevisiae regenerating protoplasts: a novel approach to study the
cell wall."
RL Yeast 15:459-472(1999).
RN [4]
RP GPI-ANCHOR.
RX MEDLINE=20469049; Pubmed=11016834;
RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
RT "Up-regulation of genes encoding glycosylphosphatidylinositol
(GPI)-attached proteins in response to cell wall damage caused by
disruption of FKS1 in Saccharomyces cerevisiae."
RL Mol. Genet. 264:64-74(2000).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
SECRETED BY REGENERATING PROTOPLASTS.

```

```

CC -1- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84162; CAA58971.1; -
DR EMBL; Z74351; CAA98873.1; -
DR EMBL; Z49209; CAA89084.1; -
DR PIR; S54039; S54039.
DR COMPLEYEST-2DPAGE; Q12355; -.
DR SGD; S0002462; PST1.
DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
DR InterPro; IPR000494; EGFR_L_domain.
DR Pfam; PF01030; Recep_L_domain; I.
KW Glycoprotein; Membrane; GPI-anchor; Signal.
FT SIGNAL 1 19
FT CHAIN 20 ?
FT PROPEP 356 444
FT DOMAIN 57 57
FT CARBOHYD 76 76
FT CARBOHYD 83 83
FT CARBOHYD 86 86
FT CARBOHYD 196 196
FT CARBOHYD 210 210
FT CARBOHYD 228 228
FT CARBOHYD 235 235
FT CARBOHYD 242 242
FT CARBOHYD 263 263
FT CARBOHYD 268 268
FT CARBOHYD 280 280
FT CARBOHYD 292 292
FT CARBOHYD 305 305
FT CARBOHYD 329 329
SQ SEQUENCE 444 AA; 45776 MW; 230F60CACA5921A4 CRC64;
Query Match 5.9%; Score 138; DB 1; Length 444;
Best Local Similarity 18.2%; Pred No. 0.98;
Matches 90; Conservative 95; Mismatches 184; Indels 126; Gaps 20;
QY 1 MHYFRNCIFFLIVILYGTNSS---PSTQNTSREWSSVOLSEESTFYLCPPVPGSTV 57
D  : : : : :
Db 3 LHSLIATALLITSALAAATSSSSIPSSCTISSHATATA---QSDLDKYSRCDTLVGNLT 59
QY 58 I--RLEFCGCMSTKKDANPNQO-----TOLEAARME-LTDLINAKAMTLASLDY 105
D  : : : : :
Db 60 IGGGLTKTALANVKE-----INGSLTIFNATNLTSPAADSLSEITDLSLQSLTILT 111
QY 106 AKIEASLSAYSEAEVTNNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPMLVEAY 165
D  : : : : :
Db 112 -----SASFGLSQSDSIKILTLPAISSFTSNKSNANNIYISDTSL-----QSVDPG 158
QY 166 KALKTTLEQRATNLEGLSS--TAYNQIRNNLVLYNKASSLTKTLDPL-----NGTLLDS 220
D  : : : : :
Db 159 SALKKVNFVNNNNKLTSLKSPVETVSDSLQSFENGQTKI--TFDDLWNANNISLTDV 216
QY 221 NEITFANKN-----INNTLSTINEQKTA-----DALNSFIKK----- 254
D  : : : : :
Db 217 HVSFANLQKINSSLGFINNSSLNFKLNTIGQTFISVNDYLNKLSFNSLSTIGGAL 276
QY 255 VQNN-----EQSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKYARRT 297
D  : : : : :
Db 277 VVANTGTLQKIGLDNLTIGTLEVGQNFSLNL--DSLKSVKCGADVSKSNFSCNA 334
QY 298 VVNGDEPSSRILANTNSITDYSWISYLAGTKYQFSNYPSTGPGSTGYLYFYKLYKAA 357
D  : : : : :
Db 335 L-----KALQKKGIGKESFVKNGASSTSVKLSSTSKSQSS-----QTTAKVSKS 380

```


Matches 107; Conservative 100; Mismatches 164; Indels 136; Gaps 28;

QY 21 SSPSTONVTSREVSVSVQSEESTFYLCPPVPGSTVIRLEFGCMSTTKKD-----ANPN 76
 DB 363 NSPKHVSXSMDEKTIE-AADKSTY---PSIADEVEDLMD-MDITFEQQCEAGNQ 417
 QY 77 GQTOLEAA-RMELTDLINAKAM--TLASLDYAKTEASLSAYSRAE-----120
 DB 418 DGLQLEKEDLMDISVIRDSPAVNDTMAVFQSPARVKGANNSIIDSQKSVIFGDEMSIDE 477
 QY 121 -----TVNNNLNATLEQKMAKTNLESAINOANT-----DKTTEDNEHPN 160
 DB 478 TQNDGTLTLPKSNVEVTTNDVTSLEQEEENASENVSMINSEHSEIDKKSF-----M 532
 QY 161 LVEAYKA-LKTTLEQRATNLE--GLSSTAYNQRNNLDVLYNKASSLITK-----207
 DB 533 LIEERAFHSSMDVAKLEDGSSKTP-----VILASQASASLATKEPSALHNSSA 584
 QY 208 TLD---PLNGGHLDSNETTANKNNITLSINEQ---KTNADALSNSFIKKVTONNEQ 261
 DB 585 TLNMSMELDNITLLKTMQITTC-----EDISMVHESIAVELNSNKEQEQGDETLLQKNDT 639
 QY 262 SFVG---TF-----TNANVQPSNYSEFA-PSADVTVPVNYKYA--RRTWNGDEPS 305
 DB 640 SNTGANFTQGHNETSQIWNVDSEAVNTSKISTYSAFNLSINQSIKRRSLNSARES 599
 QY 306 SRLANTNSITDYSWISLAGTN---TKYQSFNSYGPSTGYLYPPYKLVKAADANNYG 361
 DB 700 PRRALENS-----TMSMGOTMEALTEYR--QNKTMQTSQDSMP-----SMSLDSG 745
 QY 362 LQYKLNNGNVQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTLSVPTGE 421
 DB 746 RDLAMNTSVR-----SPHLNSSKTAAGTPSL-----MSQN-VOLPPPPSPQ 786
 QY 422 GNANKVAPMIGN-IYLSNNENNADKIP 447
 DB 787 FEMPDFDPAVNVVILTSDEPSTEQHP 813

RESULT 13
 ALS1 CANAL
 ID ALS1 CANAL STANDARD; PRT; 1260 AA.
 AC P46590;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Agglutinin-like protein 1 precursor.
 GN ALS1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=3476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11651 / B792;
 RX MEDLINE=95272392; PubMed=7752895;
 RA Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;
 RT "Candida albicans ALS1: domains related to a Saccharomyces cerevisiae
 RT sexual agglutinin separated by a repeating motif.";
 RL Mol. Microbiol. 15:39-54(1995).
 CC -|- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
 CC -|- PFM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
 CC -|- SIMILARITY: TO YEAST SAG1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L25902; AAC14649.2; -.

KW Cell adhesion; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 1260
 FT DOMAIN 433 792
 FT REPEAT 433 468
 FT 1-1.
 FT REPEAT 469 504
 FT 1-2.
 FT REPEAT 505 540
 FT 1-3.
 FT REPEAT 541 576
 FT 1-4.
 FT REPEAT 577 612
 FT 1-5.
 FT REPEAT 613 648
 FT 1-6.
 FT REPEAT 649 684
 FT 1-7.
 FT REPEAT 685 720
 FT 1-8.
 FT REPEAT 721 756
 FT 1-9.
 FT REPEAT 757 792
 FT 1-10.
 FT DOMAIN 983 1152
 FT 2 X 26 AA APPROXIMATE REPEATS.
 FT REPEAT 983 1043
 FT 2-1.
 FT REPEAT 1092 1152
 FT 2-2.
 FT DOMAIN 399 404
 FT POLY-THR.
 FT DOMAIN 408 418
 FT POLY-THR.
 FT DOMAIN 450 455
 FT POLY-THR.
 FT DOMAIN 486 491
 FT POLY-THR.
 FT DOMAIN 522 527
 FT POLY-THR.
 FT DOMAIN 558 563
 FT POLY-THR.
 FT DOMAIN 594 599
 FT POLY-THR.
 FT DOMAIN 630 635
 FT POLY-THR.
 FT DOMAIN 666 671
 FT POLY-THR.
 FT DOMAIN 702 707
 FT POLY-THR.
 FT DOMAIN 738 743
 FT POLY-THR.
 FT DOMAIN 774 779
 FT POLY-THR.
 FT DOMAIN 874 877
 FT POLY-SER.
 FT CARBOHYD 471 471
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 579 579
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 687 687
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 723 723
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 820 820
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 886 886
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 918 918
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 973 973
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1045 1045
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1068 1068
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;

Query Match 5.9%; Score 138; DB 1; Length 1260;
 Best Local Similarity 21.2%; Pred. No. 3.6;
 Matches 103; Conservative 75; Mismatches 189; Indels 118; Gaps 21;

QY 18 GTNS-----SPSTQNVTSREVSVSVQSEESTFYLCPPVPGSTVIRLEFGCMSTTKKDAN 73
 DB 748 GTDVTIIRPPNPVTVTTEYSQ---SPATTTTTPAPGGTDTVIYE--SMSSSKIST 802
 QY 74 PN-----NCQTO-----LEAARMELTDLINAKAMTWLASLDYAKIESLS 114
 DB 803 SNDITSIFSRPHRYVNSTTSDLSSTFESSMNTPTSSISDGMLLSS-----TT 851
 QY 115 AYSEAEVNNLNATLEQKMAKTNLESAINOANTDKTTFDNEHPNLVAYKALKTTLEQ 174
 DB 852 LVTESETTESICSDGKEC-----SRLSSSGVIVNPDNNESSIVTPTASTMD 903
 QY 175 RATNLEGLSSTAYNQIRNNLDVLYNKAS-SLITKTLDPNGG-----TLDSNE 222
 DB 904 SLSSDTGISATSDNVKSGSVTSTETSTIQTTPNPLSSSVTSLTQLSSIPSVSESES 963
 QY 223 ITTANKNINNTUSTINEQKTNA-----ALSNFIFKKVIONNEQSVGFTTNANVQPSNY 277
 DB 964 KVTFTSGNDGNQSGHSDQSTSTETSTIQTTPNPLSSSVTSLTQLSSIPSVSESES 1022
 QY 278 SFV--AFSADVTVPVNYKYARRVWNGDEPSSRLANTNSITDYSWISYLAGTNKYQFSF 335
 DB 1023 STTNSITDITP-----SOPDGDNDNTSS-----TNVPVPA-TSTLASSEEDNKSG 1071
 QY 336 SNYGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVEFATSTANNTANPT-----P 391
 DB 336 SNYGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVEFATSTANNTANPT-----P 391

```

Db 1072 SHESASTS-----LKPSMGENSEGL-----TTSTEIEATTSTPTEAPSP 1109
QY 392 AVDEIKVAKIVLGLRFGONTIELSVPTGEGNMKVKAPMIGNIYLSSNENNAKIPGVR 451
Db 1110 AVSGTDVTEPTDTRQPTTIS-----TTSKTNSESVATT-----QATNEN-----GCKS 1155
QY 452 PGTEL 456
Db 1156 PSTDL 1160

RESULT 14
YMF6_YEAST
ID YMF6_YEAST STANDARD; PRT; 1140 AA.
AC Q04893;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 113.1 kDa protein in PRE5-FET4 intergenic region.
GN YMR317W OR YM9924.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lyne G., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.,"
RL Nature 387:90-93(1997).
CC -!- DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54141; CAA90835.1; -
DR SGD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match 5.9%; Score 137; DB 1; Length 1140;
Best Local Similarity 21.1%; Pred. No. 3.6;
Matches 99; Conservative 76; Mismatches 188; Indels 106; Gaps 20;

QY 19 TNSPSTONVTSREVSVQSLSEESTFYLCPPVGVSTVIRLEFGCMSITTKDANPNNGQ 78
Db 412 TSSSVSEISSTKSVMSSEVSSTATSLVSSEAP--SAISSLASSRLFESSK---NTSVTS 466
QY 79 TQLEAARMELDLINAKAWTLIASQDYAKIBASISAYSEATVNNNINATLEQLKMAKT 138
Db 467 TLVATEASSVTSSLRPSSETLAS---NSIIESSLSTGYN--STVSTTSSAASSTLGSKVS 521
QY 139 NLESAINQANTDKTTFDNEHNLVEAYKALTKTLEQRATNL-----BGLSTAYNQ 189
Db 522 SSNRMATSKTSSSTSSDLKSKSVIFGNSSTVTTSPSASISLTASPLPSVWSDITSSEASS 581
QY 190 IRNNLVLDLYNKASSLITKTLPLNGGTLDDNEITANKN--INNTLSTINEQKT-----N 243
Db 582 ISSNL-----ASSAPSD-----NNSTIASASLIATKTKNSVSVSSIVSSSTTNESEN 631
QY 244 ADALNSPIKIKVIONNEQSFVGTETNANVQPSNTSFV-AFSADVTPPVNYKYARRVWNGD 302
Db 632 LATSTSLLSNKATARSILSTSNATSNASNVPTGTFFSMSSHSVITP-----GF 679

```

```

QY 303 EPSSRLANTNSITDVSIIWYSLAGTNTYQFSSNYGSPSTGYLYFPYKLVKAADANNVGL 362
Db 680 STSSASLAINSTVVS-----SSLAG-----YSFSTPESSPTTS-----TLVTSEAPSTVS- 724
QY 363 QYKLNNGNVQVEATSTSA-----NNTTANTPTPAVDEIKVAKIVLGLRFGONTIEL-SV 417
Db 725 -----SWTISAPFINNSTARSPPSTASF-----ITESTSSISSV 759
QY 418 PTGEGNMNK-----VAPMIGNIYLSSNENNAKRI---PGYRRPG 453
Db 760 PLASGDVTTSLAAHNLTTFSAPSTSSAQLVSKSTTSSSILVTPRIDRSS 808

RESULT 15
YEF3_YEAST
ID YEF3_YEAST STANDARD; PRT; 956 AA.
AC P32618;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 106.1 kDa protein in GLY1-GDAL intergenic region.
GN YEL043W OR SYGP-ORF14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.,"
RL Nature 387:78-81(1997).
RN [2]
RP FIBRONECTIN TYPE III DOMAIN.
RX MEDLINE=97148176; PubMed=8994808;
RA Bateman A., Chothia C.;
RT "Fibronectin type III domains in yeast detected by a hidden Markov
RT model.,"
RL Curr. Biol. 6:1544-1546(1996).
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18779; AAB64999.1; -
DR PIR; S30834; S30834.
DR SGD; S0000769; YEL043W.
DR InterPro; IPR003961; FN_III.
DR SMART; SM00060; FN3; 1.
KW Hypothetical protein.
FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;

```

```

Query Match 5.9%; Score 136.5; DB 1; Length 956;
Best Local Similarity 20.6%; Pred. No. 3.1;
Matches 104; Conservative 89; Mismatches 190; Indels 121; Gaps 25;

QY 20 NSSPSTONVTSREVSVQSLSEESTFYLCPPVGVSTVIRLEFGCMSITTKDANPNNGQ 79
Db 142 NEALKTRWRNTTTSSTAMQPRNSKSEPAFLPSHYSSVLSLTFSSNIT-NSATSNNG-S 199

```

```

QY 80 QLEAARMELTDLINAKANTLASLODYAKIEASISSAYSEAEVTYNNLNATLEOLKMAKTN 139
Db 200 NLPA---YTSLTTLKDLQSFSDLLKKI---LICAQEDLHVLQOQTSLLQDFQESKLE 252
QY 140 LESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATN-----LEGI-----SSTA 186
Db 253 LELELDNL---KTHWSHE---IDLKSLKSNIKSLENSKLLTDLKIEKLNKKIDKSKK 305
QY 187 YNOIRNLDVLYNKASLITKTLDPLNGGTGLDSDNEITTANKN-----INNTLSTINEQK 241
Db 306 ISKMRNDMOKWQSEDTELLSKDTIKEKYFKLLNESNASVANINKEIESLQNEISKMEESN 365
QY 242 TNADALSNEFIKKVIONNEQSFVGTFTNANVQPSNTSFVAFSADVTPVNYKYARRTVWN- 300
Db 366 KRLNASKKSLITSIV-----VNAVVE--NDKP1A-SGELSAVLKLNLDFTLEKN 411
QY 301 -----GDEPSSRILANTNSI-----TDVSW-----IYSLAGTNTKY-QFS 334
Db 412 GFLSNAGEEFLKLNADSLIKMIKOEISIDQLEANWKLQRSNLLKISALENQFEMS 471
QY 335 FSNYGPSTGYLYPPYK---LVKAADANNV-----GLQYKLN--GNVQOVEFA 377
Db 472 LNNRNLKTKLMVQPYKNNGDSLAATNSNNSAEKNRSSGSIQLPLSNMSTRGSDILIS-N 530
QY 378 TSTSANNTTA-----NP---TPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMM 425
Db 531 NKSINNSNADSAPPLRLHNPVSYSPSNEPIQPSLLSOLT--QPTDNRS----- 579
QY 426 KVAPMIGNIYLSNNENNADKIPY 449
Db 580 ----MLSN-HISSNNENKQOPSSY 598

```

Search completed: October 8, 2003, 17:03:35
Job time : 11.2108 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:00 ; Search time 15.6206 Seconds
(without alignments)
2807.386 Million cell updates/sec

Title: US-09-147-052-2
Perfect score: 2324
Sequence: 1 MHYFRRNCIEFLVILYGTN.....SSNENNADKIPGYRPGCTFL 456

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	832	35.8	702	2 S48754	major surface prot
2	754	32.4	702	2 S48753	major surface prot
3	749	32.2	650	2 S48751	major surface prot
4	745.5	32.1	649	2 S48752	major surface prot
5	715.5	30.8	647	2 A49218	hemagglutinin homo
6	578	24.9	386	2 S48755	major surface prot
7	314	13.5	865	1 VGBERR	glycoprotein b pre
8	189.5	8.2	320	2 S51560	major surface prot
9	177	7.6	6713	2 B89921	hypothetical prote
10	173.5	7.5	933	2 S41539	hypothetical prote
11	173	7.4	1302	1 J66009	fibrogen-binding
12	166.5	7.2	3890	2 C89921	surface-located me
13	166.5	7.2	4688	2 F82885	hypothetical prote
14	166	7.1	661	2 AG2422	hypothetical prote
15	165.5	7.1	2481	2 D90011	EmtB protein limpo
16	161.5	6.9	1237	2 D71850	probable outer mem
17	160.5	6.9	1072	2 A66827	hypothetical prote
18	157	6.8	820	2 T17519	cell surface antig
19	156	6.7	807	2 B71605	hypothetical prote
20	155.5	6.7	2399	2 H71879	toxin-like outer m
21	154.5	6.6	682	2 S44131	subtilisin-like pr
22	154.5	6.6	1365	2 T30822	lmp1 protein - Myc
23	152	6.5	989	2 D89832	fibrogen-binding
24	152	6.5	2660	2 E85822	probable invasin z
25	151.5	6.5	751	2 T40462	ser-lys rich hypot
26	151	6.5	624	2 PC6003	surface membrane p
27	150.5	6.5	1051	2 T18351	lmp1 protein - Myc
28	147.5	6.3	1073	2 S14032	kinesin-related pr
29	147	6.3	1487	2 AG2560	hypothetical prote

30	146.5	6.3	1085	2 T38378	kinesin-like prote
31	146	6.3	719	2 S55119	hypothetical prote
32	145	6.2	1107	2 AC0976	probable autotrans
33	145	6.2	1335	2 T17508	glycoprotein Vp260
34	144.5	6.2	796	2 T21460	hypothetical prote
35	142.5	6.1	1116	2 D97001	probable membrane
36	142	6.1	135	2 B49218	hemagglutinin homo
37	141.5	6.1	1645	2 F96907	phage-related prot
38	141.5	6.1	2401	2 T28676	rhoxy protein -
39	141	6.1	1314	1 TBYR6	transcriptional regu
40	140.5	6.0	589	2 B97806	hypothetical prote
41	140.5	6.0	3194	2 D71917	toxin-like outer m
42	140	6.0	1381	1 S45781	probable calcium-b
43	139	6.0	953	2 C89824	hypothetical prote
44	138	5.9	444	2 S54039	hypothetical prote
45	138	5.9	868	2 G71691	hypothetical prote

ALIGNMENTS

RESULT 1

S48754
major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48754
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48754
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-702 <MAR>
A:Cross-references: EMBL:L28424
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match		35.8%;	Score 832;	DB 2;	Length 702;
Best Local Similarity		40.9%;	Pred. No. 2.3e-38;		
Matches 177;		Conservative 86;	Mismatches 124;	Indels 46;	Gaps 9;
QY	50	PPPVGSTVIRLEFGCMSTIKKDNPNNG-----QTGLEAARMELTDLINAKAMTLASIQ	103		
Db	54	PNPGGMMGMNGG-----NTNPGGGTDTNAAQAAAKKELSDLLATQNSLSTVA	106		
QY	104	DYAKTEASLSAYSAAETVNNLNATLEQLMAKNLESAINQANTDKTTFDNEHPNLVE	163		
Db	107	DYANIQNTLTAAYTTAKTSNTSATLEQVKSATSLQTAIDTAASSKTSFDEKNPELIK	166		
QY	164	AYKALKTTLEORATNLEGLSTPAYQIRNNLDVLYNKASSLTKTLDPLNGTLLDSNEI	223		
Db	167	AYNALKETLKKWRNLSGLTDSNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAV	225		
QY	224	TTANKNINLTSTINEQKTNDALNSFTKKVIQNEQSFVGTFTNANVQPNYSFVAPS	283		
Db	226	SOANTNISNAYSKLETKTNTATLTSFVKVVLVKNKLTGIDT-TNNQOPQPNYSFVGS	284		
QY	284	ADVTP-----VNYKYARTVWNGD-----EPSSRILANTNSITDVSMTYISLAGNTKYQ	332		
Db	285	VDVTTGSDNARNPNWSFAQRKVTSTNTDILSQPQAEQENQSQAPDVSWIYNLTGMGAKYS	344		
QY	333	FSFSNYGPGTGYLYPYKLVRAADANNVCLQKLNNGVQVEFATS-----TSAN	383		
Db	345	LTENTYGPSTGLFYPPYKLVNSSDSQKVALEYKLNESAVKTI DFSQTSPTSPVASDATREN	404		
QY	384	N-----TTANPTPAVDIEIKVAKIVLSGLRGQNTIELSVPT-TGEGNNKNVAPMIG	432		
Db	405	NRSTAAPAGQSGTEINPAPILDDIKIAKTVLSNLKFGSNTIEFSVPTTAKEGTSKVAPMIG	464		
QY	433	NIYLSSENNAADK	445		

Query Match 32.1%; Score 745.5; DB 2; Length 649;
Best Local Similarity 40.2%; Pred. No. 1.2e-32;
Matches 177; Conservative 65; Mismatches 135; Indels 63; Gaps 12;

Qy 53 VGSTVIRLEFGCMSTKK-----DANPNNGQ-----TOLEAARMELTD 90
Db 15 IGSFVMLAAASCTTPTNPPTNPNNPPSGGMNGDTPGQGGMNNAASQELAAARMGLTT 74

Qy 91 LINAKMTLASQDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTD 150
Db 75 IFDSKAKNLGLYVDYKKTQNTLTTRAYDAAKTVDLNDSSSTQNLNEAKTRLETAIRTAATS 134

Qy 151 KTFEDNEHPNLVEAYKALKTTLEQRATNLEGLSTAYNQIRNNLVLYNKASSLITKTLTD 210
Db 135 KQTFDEQHAELVYKELKTTLSNETATLAPYADAQYAGIKMHLUSGLYDAGKAITTKTLE 194

Qy 211 PLNGGTLSDNSNEITTANKNNNTL--STINEQKTNADALSNSFIKKVIQNEQSFVGTFT 268
Db 195 PVEGDP--LTASAVNMANTKIWEAKDEVLPQKENATKLADSFVKQVLVREKINGVEAH 253

Qy 269 NANVQPSNYSFVAFSADVTVP-----NKKYARRTW--NGDEPSSRIILANT----- 312
Db 254 N--KAQPAANYSFVGYSDVTGTTGTSIPNWDYAQRTIFTNSDEP--RSISNTPADGQTM 310

Qy 313 -NSITDYSWISLAGTNTKYQFSNYGSPSGYLYFPYKLVKAADANNVGLQYKLVNNGV 371
Db 311 AQLUSNWSWISLAGTNTKYQFSNYGSPSGYLYFPYKLVNTSDQVKGLEKLVNND--- 367

Qy 372 QQVEFAFSTSA-----NNTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNK 426
Db 368 -----ARKPSAITGSDQTMNGKTPVTYNDINVAKVTLANLNFSGNKIEFSVPA-----EK 417

Qy 427 VAPMIGNIYSSNENNADKI 446
Db 418 VSPMIGNIYSSNENNANKI 437

RESULT 5
A49218
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin
A:Reference number: A49218; MUID:93162830; PMID:8432610
A:Accession: A49218
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-647 <MAR>
A:Cross-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIPI:125183)
C:Genetics:
A:Genetic code: SGC3

Query Match 30.8%; Score 715.5; DB 2; Length 647;
Best Local Similarity 39.2%; Pred. No. 5.2e-32;
Matches 174; Conservative 63; Mismatches 134; Indels 73; Gaps 13;

Qy 53 VGSTVIRLEFGCMSTKK-----DANPNNGQ-----TOLEAARMELTD 90
Db 15 IGSFVMLAAASCTTPTNPPTNPNNPPSGGMNGDTPGQGGMNNAASQELAAARMGLTT 74

Qy 91 LINAKMTLASQDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTD 150
Db 75 IFDSKAKNLGLYVDYKKTQNTLTTRAYDAAKTVDLNDSSSTQNLNEAKTRLETAIRTAATS 134

Qy 151 KTFEDNEHPNLVEAYKALKTTLEQRATNLEGLSTAYNQIRNNLVLYNKASSLITKTLTD 210
Db 135 KQTFDEQHAELVYKELKTTLSNETATLAPYADAQYAGIKMHLUSGLYDAGKAITTKTLE 194

Qy 211 PLNGGTLSDNSNEITTANKNNNTL--STINEQKTN-----ADALSNSFIKKVIQNEQSFV 264
Db 195 PVEGDP--LTASAVNMANTKIWEAKDEVLPQKENATKLADSLSSIVKRTITGVEE----- 249

Qy 265 GTFTNANVQPSNYSFVAFSADVTVP-----NKKYARRTW--NGDEPSSRIILANT----- 312
Db 250 ---AHNKAQPAANYSFVGYKRWYTELLLDKQVFPNWDYAQRTIFTNSDEP--RSISNTPAD 304

Qy 313 -NSITDYSWISLAGTNTKYQFSNYGSPSGYLYFPYKLVKAADANNVGLQYKLVN 367
Db 305 GQTMQAQPLSNWSWISLAGTNTKYQFSNYGSPSGYLYFPYKLVNTSDQVKGLEKLVN 364

Qy 368 NGNVQVQVEFAFSTSA-----NNTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
Db 365 D-----ARKPSAITGSDQTMNGKTPVTYNDINVAKVTLANLNFSGNKIEFSVPA--- 413

Qy 423 NMNKVAPMIGNIYSSNENNADKI 446
Db 414 --EKVSPMIGNIYSSNENNANKI 435

RESULT 6
S48755
major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48755
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Brownling, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface protein
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S48755
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-386 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692
A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 24.9%; Score 578; DB 2; Length 386;
Best Local Similarity 38.7%; Pred. No. 9.3e-25;
Matches 145; Conservative 51; Mismatches 123; Indels 56; Gaps 11;

Qy 53 VGSTVIRLEFGCMST-----RKDANPN----- 75
Db 15 IGSFVMLAAASCTTPTTTPNHEPKPNAPKDPKPNPGCGMGMGNGNTNPGNGGMD 74

Qy 76 NGQTOLEAARMELTDLINAKAMTASQDYAKIEASLSAYSEATVNNLNATLEQLKM 135
Db 75 NSAQOLAAAKKELSDLLATQNSLSTYADYAKIRKNDLTAAYTTAETASQQAATLEQVK 134

Qy 136 AKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSTA--YNOIRNNL 194
Db 135 AASTLQTAINTAVNEKVFDENNSELVTAYTNLKTLEGENTTAAFNDSNAYNGYGIKTHL 194

Qy 195 VDLNKAASLITKTLDPNL--GGLTSDNSNEITTANKNNNTL--STINEQKTNADALSNSF 251
Db 195 LSLYNAQKAITTSTL--LNDAGQSPKNDVVKINKEITDAINPTLLNQKAKNADMLATSF 252

Qy 252 IKKYIQQNEQSFVG---TFTNANVQPSNYSFVAFSADVTVP-----VNYKARRTWNGD 302
Db 253 TKQVL--NDAQITSGSSETSMQTPQPGNYSFVGYSDVTGTSNNARNPNWNAQKRVWDTN 311

Qy 303 EPSSRIILANTSLTDYSWISYLAGTNTKYQFSNYGSPS--TCYLYFPYKLVKAADANNV 361
Db 312 RAPLAQTEQSKLTDVSWISYLSGSGMAKTYTTFDYGASNNAYLYFPYKLVQNTD--NVG 369

Qy 362 LQYKLVNNGVQVEF 376
Db 370 LQYVLNNTTTPKLVNF 384

RESULT 7

VCGBB
glycoprotein B precursor - Marek's disease virus (strain RB1B)
C:Species: Marek's disease virus
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A32402; B32402
R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Binns, M.M.; Doel, T.; Milne, B.
J. Gen. Virol. 70, 1789-1804, 1989
A:Title: Nucleotide sequence and characterization of the Marek's disease virus homologue
A:Reference number: A32402; MUID:89293086; PMID:2544666
A:Accession: A32402
A:Molecule type: DNA
A:Residues: 1-865 <ROS>
A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837
A:Accession: B32402
A:Molecule type: protein
A:Residues: 250-271; 304-330 <ROS2>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-865/Product: glycoprotein B #status predicted <GPB>
F:709-728/Domain: transmembrane #status predicted <TN1>
F:732-752/Domain: transmembrane #status predicted <TN2>
F:727,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.5%; Score 314; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 9.1e-10;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRNCIFFLIVLYGTNSPSPQNTSREVSVSVQVLSSESTFYLCPPVPGVSTVIRL 60
Db 1 MHYFRNCIFFLIVLYGTNSPSPQNTSREVSVSVQVLSSESTFYLCPPVPGVSTVIRL 60

QY 61 E 61

Db 61 E 61

RESULT 8

SS1560
major surface protein (clone pMGAL.6) - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C>Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S51560; S48757
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surface P
A:Reference number: S48751; MUID:95010739; PMID:7925999
A:Accession: S51560
A:Molecule type: DNA
A:Status: nucleic acid sequence not shown
A:Residues: 1-320 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAAG2415.1; PID:g535688
C:Genetics:
A:Genetic code: SGC3

Query Match 8.2%; Score 189.5; DB 2; Length 320;
Best Local Similarity 48.9%; Pred. No. 0.0016;
Matches 46; Conservative 11; Mismatches 26; Indels 11; Gaps 3;
QY 354 AADANNVGLQKLNNGNQVVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTI 413
Db 41 SADS-----PTNQENSQSQAPASA-----MNETPTDGINVAKVTLTLKFGSNTI 90
QY 414 ELSVP-TGEGNNKVPAMIGNIYLSNENNADKI 446
Db 91 ELSVPTTDECTSKVAPMIGNIYTSNDQNRKI 124

RESULT 9

B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BABA2527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Query Match 7.6%; Score 177; DB 2; Length 6713;
Best Local Similarity 21.3%; Pred. No. 0.51;
Matches 105; Conservative 76; Mismatches 191; Indels 122; Gaps 19;

QY 20 NSSPTQNTSREVSVSVQVLSSESTFYLCPPVPGVSTVIRLEFGCMSTTKDANPNNGOT 79
Db 465 NNLTSINNAQRDLTKI-----DQAT-----TVAGVEAVSNTG-T 499
QY 80 QLEAARMELDLINAKMTLASLDYAKIEASLSAYSEAE-----VNNLNLAIEQLK 134
Db 500 QLTAMANLQNGINDKANTLAS-ENYHDADSDKKTAYTQAVTNAENILKNSSGSLD--- 555
QY 135 MAKTNLESAIQANTDKTTFDNEHPNLYEAYKALKITLEQATNLEGLSSTAYNQIRNML 194
Db 556 --KAAVENALSOVTKAGALNGNH-NLEQAKSNANTTI-----NGLQLHTLAQKDLKQOV 608
QY 195 VDLYNKAS-SLITKTLDPLNG--GTLSDNETTKANKNINNTLSINEOKT----- 242
Db 609 QQAQNVAGVDVTKSSANTLNGAMGTLRNSIQDNTATKNGQVLDATERKNTYNNVADSA 668
QY 243 -----NADALNSFIKKVIONNEOSFVGTFTTNANVPNSYFVAFSADVTPVNYK 292
Db 669 NGVINATSNPNMDANAIQIATQVTSKTNALDGNLHQAQKT-----ATNAIDGATNLN 723
QY 293 YARRTVWNGDESSRILANTNSITDVSIIYSLA-----GTNTKYQFSFSNYGP 340
Db 724 KAQKDALKAQVTSQAVANVTSIQOTANELNTAMGOLQHGIDDENATKQTKYRDAEQSK 783
QY 341 STGYLYFPYKLVKAADA-----NNVGLQYKLN-----NGNVQVVEFATSPS 381
Db 784 KTAY-----DQAVAAAKAILNKQTSNSDKAAVDRLAQVTSKDALNGDAKLAERAKAR 839
QY 382 AN-----NTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNM-----KV 427
Db 840 QNLGTLNHTNAQRTALEGQINQATV-----DGVNTVKTNTANTLDGAMNSLQGAINDKD 894
QY 428 APMIGNIYLSNEN 441
Db 895 ATLQNQNYLDADES 908

RESULT 10
S41539
fibrinogen-binding protein - Staphylococcus aureus
N:Alternate names: clumping factor
C:Species: Staphylococcus aureus
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S41539; S36630
R:McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of S
A:Reference number: S41539; MUID:94224142; PMID:8170386
A:Accession: S41539
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-933 <MCD>
A:Cross-references: EMBL:Z11852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match 7.5%; Score 173.5; DB 2; Length 933;
Best Local Similarity 22.5%; Pred. No. 0.053;
Matches 111; Conservative 66; Mismatches 185; Indels 131; Gaps 23;

QY 2 HYFRNCIFFLIVL-----YGTNSS-----PSTQNTVSREVVSQVLSSEESTFYLCPPP 52
DB 9 HAIRKKSIGVASVLVGLTGFLSSKEADASESVTQSDASNESKNSDSSVSAAPKT 68
QY 53 VGSVIRLEFCGMSITKKDANPNNGQTOL--EARMELTDLINAKAMTSLASLDYAKIEA 110
DB 69 DDTNV-----SDTKTSSNTNGETSVAGNPAQQTQSSSTNAT-- 108
QY 111 SLSSAYSEAEFVNNLNATLEQLKMAKTNLESAINQAQNTDKTTFDNEHPNLVEAYKALKT 170
DB 109 EETPVTGEATTTTNOANTPATTQSSNTNABELVNO--TSNETTFND--TNTVSSVNS-- 162
QY 171 TLEQATNLEGLST-----AYNQIRNLDVLYNK-----ASSLIT 206
DB 163 --PQNSTNAEVSTQDTSTEATPSNESAPQSDASKNGVYVNOAVNTSAPMRFAFLAA 220
QY 207 KTLPLNGGTLDDNEITANKNNLTSTINEQ----KTN-ADALNSFTK----KVIQ 257
DB 221 VAADAPAGTDI--TNQLTNTVTGIDSGTIVYVPHQAGYVKLYNGFSVPNSAVKGTFTKITV 279
QY 258 NNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVMNGDEPSSRILANTNSITD 317
DB 280 PKELNLNGVTSTAKVPP-----IMAGDQ-----VLANGVIDSD 312
QY 318 VSWIYSLAG--TNTKYQSFSGNYGPGSTGYLYPPYKLVKAADANNVGLQKLNNGVQVVF 376
DB 313 GNVIYTFDYNTKDDVKATLTMPA--YI-----DPEN-----KRTGNV----- 350
QY 377 ATSTSANNTANPAPVADIEIKVAKI-----VLSGLRFGQNTIELSVPTGEGNNKVAP 429
DB 351 TLATGIGSTANKVLVDYKYGKFNLSINGTIDQIDKNTNTYRQIYYNPSGDNVIAP 410
QY 430 MI-GNIYLSNEN 441
DB 411 VLTGNLKPNTDSN 423

RESULT 11
JC6009
surface-located membrane protein lmp3 precursor - Mycoplasma hominis
C:Species: Mycoplasma hominis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: JC6009
R:Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.
J. Bacteriol. 178, 2775-2784, 1996
F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <SIG>
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene system
A:Reference number: JC6009; MUID:96213016; PMID:8631664
A:Accession: JC6009
A:Molecule type: DNA
A:Residues: 1-1302 <LAD>
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336
C:Genetics:
A:Gene: lmp3
A:Superfamily: SGC3
C:Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology
C:Keywords: duplication; membrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
F:993-992/Domain: tetratricopeptide repeat homology <TT1>
F:993-1026/Domain: tetratricopeptide repeat homology <TT2>
F:1089-1120/Domain: tetratricopeptide repeat homology <TT3>
F:1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match 7.4%; Score 173; DB 1; Length 1302;
Best Local Similarity 22.8%; Pred. No. 0.09;
Matches 94; Conservative 67; Mismatches 163; Indels 88; Gaps 18;

QY 67 ITKDDANPNQ-----TOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 122
DB 864 ITKLETFNKDKDVKFELEQTRKDIDEFIN-----TNKTNPDYSTLISELTSKRDSKNSI 919
QY 123 NNNLNATLEQLKMAKTNLESAINQAQNTDKTTFDN-----EHPN-----LVEAYKAL 168
DB 920 TNSNKS--DIETANTEQLKALAKANTDQADNLARSTKQOLNKSISANTLLAKLFDK 977
QY 169 KTTLEQATNLEGLSSSTAYNOI--RNNLDVLYNKASSL-----ITKTDPLNGGTLDDSN 221
DB 978 DNTIOQAKTELEKEVQKANQAVASNTASQMSAKSSLDKAVTEITKKLETNPKDKDVF 1037
QY 222 BITTANKNINTLSTINEQKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVA 281
DB 1038 ELEQTRKDIDEFIN--NKTNPNTSTLISELTSK--RDSKNSITSSNKSDETANTEL-- 1092
QY 282 FSADVTPVNYKYARRTVMNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFSPNYP 341
DB 1093 -----KQALAKAN--TDKAAADNLA--RSTKQOLNKSISAN 1125
QY 342 TGYLYFPYKLVKAADANNVGLQKLNNGVQVVFATSTSANNTANPTPAVDEI--KVA 399
DB 1126 T-----LLAKLTDKNTIOQAKTE--LEKEVQKANQAVASNTVSMQSAKSSLDTKVT 1176
QY 400 KIVLSGLRFGQNTIELSVPTGEGNNKVAPMIGNI--YLSNENNADKIPY 449
DB 1177 BITKLETFNKDK-----EAKFNELKTRGQIQEFINTNKN--PNY 1216

RESULT 12
C89921
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaiko, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: C89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <KUR>
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BAB42528.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhB

Query Match 7.2%; Score 166.5; DB 2; Length 3890;
Best Local Similarity 19.7%; Pred. No. 0.92;
Matches 100; Conservative 89; Mismatches 210; Indels 109; Gaps 18;

QY 17 YGTNSSPST-----QNTSREVSVSSQVLSSEESTFYLCPPVPGSTV 57
DB 2427 HVSNSPSTLPAPAAHTVNTTEIVKDYGSNTVAEINNNAVQANKRT-----ATI 2476
QY 58 -----IRLEFGCMSITKKDANPNNGOT-----QLEARMELTDLINAKA 96
DB 2477 KNGTAMPTNLGGSTTTIPVTVTYNDGSTEEVQESIFTKADKRELITARNLDDPVSPEG 2536
QY 97 MTLASLDY-----AKIEASLSAYSEAEFVNNLNATLEQLKMAKTNLESAINQAQNTDKT 152
DB 2537 KKPGTITQYNNAMNAQOQINTAKTEAOQVINNERATPQQVSDALTKKVRAAQTKIDQAKA 2596
QY 153 TFDNEHPN--LVEAYKALKTTLEQATNLEGLSSSTAYNOIRNNLDVLYN-----KASSLIT 206
DB 2597 LLQNKEDNSQLVTSKNNLQSSVNO-----VPSTA--GMTQOSIDNTNAKKREATEIT 2647
QY 207 KTLPLNGGTLDDSNITANKNINTLSTINEQKTNADALSNSFIKKVIONNEQSFVGT 266

Db 2648 AARVIDNGD-ATAQIQISDEKHRVDNALTALNOAKHDLTADTHALEQAOQLNR---TGT 2703
QY 267 FTN---ANVQPSNYSFVAFSADVT-----PV-NYKARTVWNGDEPSSRILA 310
Db 2704 TTGKKPASITAYNNSIRALQSOLTSAKNSANALIQKPIRTVOEQSALTNNVRNRLTQ 2763
QY 311 NTSNITDVSWIYSLAGTNTK---YQFSNYPSTGYLYFFPKLYKADANNVGLQYKL 366
Db 2764 AINQLVPLADNSALRTAKTLDEEINKSVTTDGMTQSSQAYENAKRAGQTTETNAQNYI 2823
QY 367 NNGNVQOQVEFATS-----TSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSV 418
Db 2824 NNGDATDQIAAEKTKVEKYNSLQKAIAGLPDLAPLQATQTL-----QNDID--QP 2875
QY 419 TEGGNKVKVPMIGNIYLSNENNADKI 446
Db 2876 TSTTGMTSASVAFAFNDKLSAARTKIQEI 2903

RESULT 13
F82885
hypothetical protein U0482 [Imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi
A:Reference number: A82870
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4688 <GLA>
A:Cross-references: GB:AE002145; GB:AF222894; NID:g6899476; PIDN:AAF30894.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0482
A:Genetic code: SGC3

Query Match 7.2%; Score 166.5; DB 2; Length 4688;
Best Local Similarity 21.1%; Pred. No. 1.2;
Matches 112; Conservative 78; Mismatches 166; Indels 175; Gaps 25;

QY 19 TNSPSTQNTVSRVSVSSVOLSEESTF-----YLCPPPVGSGTVIRLEGCMXI 67
Db 1387 SNNVSLKANTWIEITKRLVNVNDQSLISAKIEVDIDNVLTADKPNIVIOLE----- 1440
QY 68 TKKDANPN-----NGQTOLEAARMELT-DLINAKAMTLASLDQYAKIEASLSS----- 114
Db 1441 NSNDANNLKLATNAQVVVNNKKFKFDLVSLKIN-----QNYVKEISFNSKPTNAVY 1495
QY 115 -----AYSEAE--TVNNNLNAT-----LBQMKAKTN 139
Db 1496 NPTNKNNTNIVSYDEQNKISLNINPTSYKPTINKNDSVNDVDLQVDKQLLANQY 1555
QY 140 LESALNQAQNTDKTT-----FDNEHPNLVEAYKALKTTLEORATNLEGLSSPAYQIRNN 193
Db 1556 LRLKQLNDKNTVWTDPIFN-----NAKISEK-LSNLHNRAYELEGL---YFDDQNS 1608
QY 194 LVDLYNKASSLITKTLDPINGTGLDSNEITTANKNINNTLSTIN-----EOKTNAD 245
Db 1609 VNDMTNNOISFNSKIHKP-----KIEFEPSLTINYDTNNAIKTVSAHNAHVHFKLKTND 1664
QY 246 ALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTTPVNYKARTVWNGDEPS 305
Db 1665 ALENDQIVIEAV-----FAPTNLNDQKV-----E 1689
QY 306 SRLANTNSITDVSWIYSLAG--TNTKYOF---SFSNPGSTGYLYFPYKLVKAADANNV 360
Db 1690 AKLNVTNFSNFEGLFNLGLRRETTYRLIKVFKNK-----PNKAYELLNKNV 1740
QY 361 GLQYKLNNGNVQOQVEFAT-----STSANNTTANPTPAVDEI-----KVAKIV 402

Db 1741 IFEYK--NGS-QAYEFTTQKFEHKVIDVSVSSSTNTTQOEITVKIDGIORANNKKLELV 1797
QY 403 LSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNENNADKIPCYRRPG 453
Db 1798 YESNIIIGDEIKTV---DNNNN-----SVHLSFDKKEYNLVNLNLPK 1838

RESULT 14
AG2422
hypothetical protein all4935 [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG2422
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-661 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB76634.1; PID:g17134073; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4935

Query Match 7.1%; Score 166; DB 2; Length 661;
Best Local Similarity 22.4%; Pred. No. 0.086;
Matches 106; Conservative 50; Mismatches 216; Indels 102; Gaps 17;

QY 1 MHYFRNCIFFLI--VILYGTNSS-----PSTQNTVSRVSVVOLSEESTFYLCPPPV 53
Db 1 MNLVTQKALIGLITTAALGTILASIKPATATQTITISAN-VKPIQVAIKEA-----PEAI 54
QY 54 GSTVIRLEGCMSTYKKNPNNGQTOLEAARMELTDLI-----NAKAMTLASLDY 105
Db 55 -----KK-----AQEAADALKKAEDDLSGLIRSTRNRAKSDANSRLNRAEQDF 96
QY 106 AKIEASLSAYSEAEVTNNLNATLEQLKWKATNLESALNQAQNTDKTTFDNEHPNLVEAY 165
Db 97 TQADAGFQTAQTALNTATNNRNNAAQNALNTATNNRNNAAQNALNTATNNRDN-----AO 149
QY 166 KALKTTLEQATNLEGLSTAYNQIRNNLVLYNKASSLITKTLDPINGTGLDSNEITT 225
Db 150 NALNTATNNR--NNAQNALNTATNN--RNAQNALNTATNNRNNAAQNAL-----NTATN 199
QY 226 ANKNINNTLSTINEOKTNADALSNSFIKKVIQ---NNEQSFVGTFTNANVQPSNYSFVAF 282
Db 200 NRNAQNALNTATNNRNNAAQAEVDATRNLAQARRGNSQAKIQAINALNQAQNTLRNTAN 259
QY 283 SADVTTPVNYKVARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSNYPST 342
Db 260 TALNTATNNFNTANTAFN-----TANNFST-----ANTAFNTATNNFSTA- 300
QY 343 GYLYPPYKLVKAADANNVGLQYKLNNGVQOQVEFATSTSANNTTANPTPAVDEIKVAKIV 402
Db 301 -----NTALNTATNNFNTATIAELDQANTRLNTARN-----DFNTANSN 338
QY 403 LSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNENNADKIPGYRRPGTFL 456
Db 339 FSRTGNELNTATNNFNTANTNTATNFNN--ASSRRNTAEQARNQVREETRL 390

RESULT 15
D90011
FntB protein [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: D90011
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2481 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701961; PIDN:BAB43253.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fmb(mrp)

```
Query Match          7.1%; Score 165.5; DB 2; Length 2481;
Best Local Similarity 21.9%; Pred. No. 0.56;
Matches 104; Conservative 66; Mismatches 173; Indels 131; Gaps 22;

QY  25  TQNTSREVSVSVOLSESESTFYLCPPVGVSTVIRLEFGCMSITKK---DANPNNGOTQL 81
DB  1150  TQDITAEIATAA-----NADVDNAVTVQANSNIEAANSQNDVDQAKTTGETSI 1196

QY  82  E-----AARMLTDLINAKAMTLASLDYAKIEASLSAYSEAEVNNLN----- 127
DB  1197  DQVTPVNVKATARNEITAILNNKLOEIQATPDATDEEQAADA--EANTENGRANQAIS 1254

QY  128  --ATLEQLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTL---EQRATNLEGL 182
DB  1255  AATTNAQVDEAKANAARAIN-AVTPKVYKQAQKDEIDQLQATQTNVINNDQATNEEKE 1313

QY  183  S-----STAYNQIRNNL-----VDLYNKASSLITKTLDPNGGTLDSN-----E 222
DB  1314  AAIQQLATAVTDAKNNITATDDNGVDPAKDAGKNSIQSTQP---ATAVKSNAKNEVDQA 1370

QY  223  ITTANKNINNTLSTINEQKTNADALSNSFIKKVIQNNQSEFVGFTTNANVQPSNYSFYAF 282
DB  1371  VTTQQAIDNTTGATTEKNAKDL-----VLKAKEKAYQDIL---NAQTTN----- 1414

QY  283  SADVTPVNYKYARRTVHNGDEPSSRILANTNSITDVSWIYSLA----GTNTKYQFSFSNY 338
DB  1415  --DVTQI-----KDQAVADIQGITADTTIKDVAKDELATKANEOKAL--- 1454

QY  339  GPSTGYLYFPYKLVKAADAN-----NVGLQYKLNNGVQVQVEFATSTSANNTTA-NPT 390
DB  1455  -----IAQTADATTEEKQANQVDAQLTQGN-QNIENAQSIDVDVNTAKDNAI 1501

QY  391  PAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVKAPMIGNIYLSNENNAD 444
DB  1502  QAIDPIQASTDVKTNAR-----AELLTEM---QNKITEILNNTTNEEKGN 1546
```

Search completed: October 8, 2003, 17:08:23
Job time : 18.6206 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	156	6.7	807	5	US-09-820-843B-108	Sequence 108, App
2	142.5	6.1	770	5	US-09-897-516A-4453	Sequence 4453, App
3	138	5.9	1010	6	US-10-415-657-10	Sequence 10, Appl
4	134	5.8	861	5	US-09-820-843B-109	Sequence 109, Appl
5	131	5.6	523	6	US-10-425-114A-58966	Sequence 58966, A
6	124.5	5.4	1225	5	US-09-581-286A-485	Sequence 485, App
7	124.5	5.4	1226	5	US-09-581-286A-484	Sequence 484, App
8	124.5	5.4	1251	5	US-09-581-286A-351	Sequence 351, App
9	124	5.3	615	6	US-10-650-274-166	Sequence 166, App
10	123	5.3	930	5	US-09-200-650E-3	Sequence 3, Appl
11	123	5.3	3241	6	US-10-647-057-1	Sequence 1, Appl
12	122	5.2	1531	1	PCF-US02-37235-44	Sequence 44, Appl
13	121	5.2	945	1	PCF-US03-20460-14	Sequence 14, Appl
14	120	5.2	1315	5	US-09-200-650E-5	Sequence 5, Appl
15	119.5	5.1	1386	6	US-10-472-078-8	Sequence 8, Appl
16	119.5	5.1	1449	6	US-10-472-078-2	Sequence 2, Appl
17	118.5	5.1	1033	5	US-09-820-843B-75	Sequence 75, Appl
18	118	5.1	386	1	PCF-US03-20322-229	Sequence 229, App
19	118	5.1	1881	1	PCF-US03-27401-316	Sequence 316, App
20	117.5	5.1	477	6	US-10-425-114A-71726	Sequence 71726, A
21	117.5	5.1	1222	6	US-10-650-274-142	Sequence 142, App
22	117	5.0	918	5	US-09-200-650E-1	Sequence 1, Appl
23	116	5.0	741	6	US-10-425-114A-64315	Sequence 64315, A
24	115.5	5.0	1032	1	PCF-US03-20460-4	Sequence 4, Appl
25	114.5	4.9	990	6	US-10-047-676B-7	Sequence 7, Appl
26	114	4.9	402	1	PCF-US02-34769-6	Sequence 6, Appl

```
Db      746 NMLTTSN 752
      1 11:1
      746 NMLTTSN 752

RESULT 2
US-09-897-516A-4453
; Sequence 4453, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 4453
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-4453

Query Match      6.1%; Score 142.5; DB 5; Length 770;
Best Local Similarity 21.2%; Pred. No. 1.4;
Matches 88; Conservative 70; Mismatches 177; Indels 81; Gaps 19;

QY 23 PSTQNTSVREWSVOLSEESTFY-----LCPPVPGSTVIRLEFGCMSITKKDANPN 76
Db 405 PLTRKINHKLVSIDINLAADVAYNRTETDALIPIQDI-----AQAMKLADTAQ 458
QY 77 GQTOLE--AARMELTDLINAKMTLASLDYAKIEASLSAYSEAEVTNNLNATLEQLK 134
Db 459 GATAIRDAASKVPLTRRNGK-----ELVNDIKLIASDVNAY--DKBEIDQIDGVKELAN 512
QY 135 MAKTNLESAINQANT--DKTFDNEHPNLVEAYKALKTLTEQATNLEGLSSYAYNQIRN 192
Db 513 AANNADSKVPVPTINNKKALLDMLNASDVDTYAKGEIDQINSVRKLANDANNVNG 572
QY 193 N--LVLDYNNKASLLITKTLPLNGTGTLDSNEITT-----ANKVINNTLSTINDQKTNADA 246
Db 573 KVELTRTVNNKALLDTITLAFDVG--YNNKNEIDSRLDKVTKNANGRLA---KDENGADI 628
QY 247 L-SNSFIKKVIONNEQSFVGTFTTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPS 305
Db 629 PDRNAPVKNI-----GLGDLIGSKIESQ-----LIGQDAT----- 658
QY 306 SRLANTNNTSDVSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQYK 365
Db 659 ---IINLGKITQISGV-AIAGTIK-QENTSIYGVYTYNY-YK-----IRLPVS 703
QY 366 LNNGVQVVEFATSTSANNTTANPTPAVDIKVAKIVLSGLRGQNTIELSVPTGE 421
Db 704 LPNG-----IISCHASIACNPNDSOLPSHLADVQRQNSDGVLSQDTLTISVTTPE 756

RESULT 3
US-10-415-657-10
; Sequence 10, Application US/10415657
; GENERAL INFORMATION:
; APPLICANT: Genix Bioscience GmbH
; TITLE OF INVENTION: Eukaryotic cell division genes and their use in diagnosis and tre
; TITLE OF INVENTION: of proliferative diseases
; FILE REFERENCE: CB617730S
; CURRENT APPLICATION NUMBER: US/10/415,657
; CURRENT FILING DATE: 2003-09-08
; PRIOR APPLICATION NUMBER: US 60/246,750
```

```
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: C. elegans
US-10-415-657-10

Query Match      5.9%; Score 138; DB 6; Length 1010;
Best Local Similarity 21.1%; Pred. No. 2.8;
Matches 107; Conservative 100; Mismatches 164; Indels 136; Gaps 28;

QY 21 SSPSTQNTSVREWSVOLSEESTFYLCPPVPGSTVIRLEFGCMSITKKD-----ANPN 76
Db 363 NSPKHVKSKSMDEKTIIE-AADKSTRY---PSTADEVEDLDMD-MDITEQPCPCAGNOQN 417
QY 77 GQTOLEAA--RMELTDLINAKM--TLASLDQYAKIEASLSAYSEAE----- 120
Db 418 DGLQKQEDLMDLSVIRDSFPAVNDTMAVQSPARVKGANNISLDSQKSVIFGDEMISIDE 477
QY 121 -----TVNNLNATLEQLKMAKTNLESAINQANT-----DKTTFDNEHPN 160
Db 478 TQNDGTLTLPKSNVEVTTTNDVYLSLREQEENASENVSMINESSVHSEIDKKSF-----M 532
QY 161 LYEAYKA-LKTTLEQATNLE--GLSSTAYNQIRNNLVLYNKASSLITK----- 207
Db 533 LIEERAFMHSSMIDVAQKLEDDGSSKTP-----VILASQSASLATKEPSALHNSA 584
QY 208 TLID---PLNGTGTLDSNEITTANKNINNTLSTINEQ---KTNADALSNSFIKKVIONNEQ 261
Db 585 TLNNSMELDNNTLLKTMQITTC-----EDISWVHESIAVELNSNKEQEQFGDETLOKNDT 639
QY 262 SFVG-----TF-----TNANVQPSNYFVA--FSADVTPVNYKYA--RRTVWNGDEPS 305
Db 640 SNTGANFTFOGHNETSQIMNNDSEAVNTSKISTYSAFNLSINQISKRRLSLSARES 699
QY 306 SRLANTNNTSDVSWIYSLAGTN-----TKYQFSNYPSTGYLYFPYKLVKAADANNVNG 361
Db 700 PRVALENS-----IMSNQGTMEALTEYR---QNKTMQTSQDSMP-----SMSLNDSG 745
QY 362 LQYKLANGNVQVVEFATSTSANNTTANPTPAVDIKVAKIVLSGLRGQNTIELSVPTGE 421
Db 746 ROLANTNTSVR---SPLNNSKTAAPGTFSL-----MSQN-VOLPPPSFQ 786
QY 422 GNMNKYAPMIGN-IYLSNENNNADKIP 447
Db 787 FEMPDPDPAVNVVYLTSEDPSREQHP 813

RESULT 4
US-09-820-843B-109
; Sequence 109, Application US/09820843B
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
; TITLE OF INVENTION: PROTEINS USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843B
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 109
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: hypothetical protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: gi|4493994
US-09-820-843B-109
```

```

QY 130 LEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQATNLE-----180
Db 346 -QKMQQOQLEQKLKELESFRSQEBHEKRMQAESALLSQGKELAQSHSEEVRLTIEI 404
QY 181 GLSTAYNQIRNNLVLY-----NKASLTKTTL-DPLNGTGLDS-NE 222
Db 405 KMANEKINELKQTKEDLHDVTCELKRDVERITEQNSQSEVLRELIGDEIN--TLKDSKNE 462
QY 223 ITTANKNINNTLSTINEQKTNADALS N 249
Db 463 LQSEIKSLKSTISQINTYK-NAAALQH 488

RESULT 6
US-09-581-286A-485
; Sequence 485, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHSEL, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09581,286A
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 485
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (982)..(982)
; OTHER INFORMATION: xaa can be any naturally occurring amino acid
US-09-581-286A-485

Query Match 5.4%; Score 124.5; DB 5; Length 1225;
Best Local Similarity 21.6%; Pred. No. 15;
Matches 98; Conservative 51; Mismatches 155; Indels 149; Gaps 22

QY 113 SSAYSEATVNN-----LNATLEQLKMAKTNLESAINQANTDKTTFDNEHP-NLVEAY 165
Db 358 NAAINEAQFAKNSGYTHICGYLDGDFALANNLSKLTA---TDENHFFATPANAFAAF 413
QY 166 KALKTTLEQATNLEGLSSATYQNIRNNLVLYNKASSLITKTLPLNG-----214

```

Db 414 DNIAQTINIGRGEVTFVAPGFIVKNLQ-----SGDVTHLLNVSGTVHYDVSTKKL 468
QY 215 ---GTLDSNEITTA-----NKNIN-NFTLSTINEQKTNADALSNSFIKKVION 258
Db 459 TWTGTLISSEATYRIYADLDYIQNDIPVNTTSAIGDPLGGFD-----N 517
QY 259 NEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARFTVWNGDEPSSRILANTNSITDV 318
Db 518 TEAKL--TYTNSNGEPNQ--LIFPRPTVKLGYGVIKRYHVLVKNKDGOPQIANGTVVSSL 573
QY 319 S-----WT-----YSLAGTNT-----KYOF----- 333
Db 574 SEAHVLSQDFELPSGGGHIYVKKIKDKTEALQYYSVPPTNTVITTTADGKRYRFEVVP 633
QY 334 -SFSNYG-----PSTGYLYFPYKLVK-----AADANNVGLQYKLNNGVQ 372
Db 634 GSTPNPGGIGISWKKPAGNAIFAYKLLNVWGGTTDQOSEWDTVSNWTGAQVPLTG---E 690
QY 373 QVEFATSTANNTANPTPAYDEIKV----AKIVLSGLRFGONTIELSVPTG-----EGN 423
Db 691 DVEFATTENFG-----SPVADLHVPTTNPKEII--GNLNNSDKDLVVTSSQLTINGV 742
QY 424 MNKVAPMIGNIYLSNENNADKI-----PGYRR 451
Db 743 VEDNNPNVGTIVVKSNDNPTGTLFANPGYNQ 775

RESULT 7

US-09-581-286A-484
; Sequence 484, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHIEL, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 484
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (983)..(983)

; OTHER INFORMATION: xaa can be any naturally occurring amino acid
US-09-581-286A-484

Query Match 5.4%; Score 124.5; DB 5; Length 1226;
Best Local Similarity 21.6%; Pred. NO. 15;
Matches 98; Conservative 51; Mismatches 155; Indels 149; Gaps 22;

QY 113 SSASEATVNN-----LNATLEQLKMAKTNLESAINQANTDKTTTONEHP-NLVEAY 165
Db 359 NAAINEAQFAKNSGYTIHTIGYDLGDFALANNLSKLT-----TDENHFFATPANLAAAF 414
QY 166 KALKTTLEQRATNLEGLSLSTAYNQIRNNLVLDLYNKASSLITKTLDPANG----- 214
Db 415 DNIAQTINIGRGEVTFVAPGFIVKNLQ-----SGDVTHLLNVSGTVHYDVSTKKL 469
QY 215 ---GTLDSNEITTA-----NKNIN-NFTLSTINEQKTNADALSNSFIKKVION 258
Db 470 TWTGTLISSEATYRIYADLDYIQNDIPVNTTSAIGDPLGGFD-----N 518
QY 259 NEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARFTVWNGDEPSSRILANTNSITDV 318
Db 519 TEAKL--TYTNSNGEPNQ--LIFPRPTVKLGYGVIKRYHVLVKNKDGOPQIANGTVVSSL 574
QY 319 S-----WT-----YSLAGTNT-----KYOF----- 333
Db 575 SEAHVLSQDFELPSGGGHIYVKKIKDKTEALQYYSVPPTNTVITTTADGKRYRFEVVP 634
QY 334 -SFSNYG-----PSTGYLYFPYKLVK-----AADANNVGLQYKLNNGVQ 372
Db 635 GSTPNPGGIGISWKKPAGNAIFAYKLLNVWGGTTDQOSEWDTVSNWTGAQVPLTG---E 691
QY 373 QVEFATSTANNTANPTPAYDEIKV----AKIVLSGLRFGONTIELSVPTG-----EGN 423
Db 692 DVEFATTENFG-----SPVADLHVPTTNPKEII--GNLNNSDKDLVVTSSQLTINGV 743
QY 424 MNKVAPMIGNIYLSNENNADKI-----PGYRR 451
Db 744 VEDNNPNVGTIVVKSNDNPTGTLFANPGYNQ 776

RESULT 8

US-09-581-286A-351
; Sequence 351, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTHIEL, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654


```

US-10-650-274-166

Query Match          5.3%; Score 124; DB 6; Length 615;
Best Local Similarity 22.9%; Pred. No. 9.1;
Matches 97; Conservative 46; Mismatches 159; Indels 122; Gaps 20;

Qy 19 TNSPSTQNTSREVSSVOLSEESTFYLCPPPVGVSTVIRLFGCMSITKKDANPNNGQ 78
    :||| ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 46 SNSSSSTAN-----KSKVF-----PVSYNNTAKAIGGNVNAVVDNSPPKG- 87

Qy 79 TQLEAARMELTDLINAKAWTLASLODYAKIEASLSAYSEATVNNNLNATLEOLKMAKT 138
    : : : : : ||| ||| ||| ||| :||| :||| :||| :||| :||| :||| :|||
Db 88 -----VFNEELYTDAYDNDYMSPAESLFAYNSTFFKFNNGAATIKODNSAKT 135

Qy 139 NLESAINQANTDKTTFNEHP-----NLVEAYK--ALKTTLEQRAT-----NLEGLSSTAYN 188
    : : : : : ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :||| :|||
Db 136 -----IVTTIKPNVKWSDGQPVVARDLVYAYEIMANKATKSQRTESLQNI EGL--TEYH 188

Qy 189 QIRNLVDLYNKASSLITKTLDPUNGGTGLDLSNEITTFANKNNINWLTSTINEQKTNADALS 248
    || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 189 D-----GKADTISGLTMPKGDNG-----NTMVIHFQKMPFSNTSGNGYLESAAP 234

Qy 249 NSFIKKVIQNEQS-----FVGFFTNANQPSNYSVFAESADTVPNYKRYARTVW 299
    || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 235 YHYLHDVAFKLEKSDSKVRKOPLFQFGYKISKVVAGQ-----SVEYTPNQY-----YW 282

Qy 300 NGDEPSSRI--LANTINSITDYSWIYSLAGTNTKYQF-----SFSNYG-PS 341
    || :||| :||| :||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 KGRPSLSKITFENVSSASITS-----ALKNHKYDIVGMPSDSYDMKNISGYTNLGHQS 337

Qy 342 TGYLYFPYKLVKAAD--ANNV-----GLOYKLNNGNVOOVFEFATSTSANNT 385
    || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 338 LAVNYLFGKLGKWDKKSENVDPNSKMANKSLRQAMGYALNNDQV--AAKPYNCTRSRAT 396

Qy 386 TANP 389
    | |
Db 397 TLIP 400

RESULT 10
US-09-200-650E-3
; Sequence 3, Application US/092006050E
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eldhinn, Delidre N1
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200.650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-3

Query Match          5.3%; Score 123; DB 5; Length 930;
Best Local Similarity 20.9%; Pred. No. 14;
Matches 90; Conservative 60; Mismatches 167; Indels 114; Gaps 18;

Qy 53 VGSVIRLEFCMCSTTKDANPNNGQQTQLEARMELT-----DLINAKAWTLAS 101
    ||| :||| ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 35 VGTTLI---FGLSGHEAKAAEHTNG--ELNOSKNETAPSENKTKTKVDSQLKDNQTOTA 89

Ov 102 LODYAKIEASLSAYSE--AETVNNNLNATLEOLKMAKTNLESAINQANTDKTTFDNE--H 158

```

Db 90 TADQPKVTHSDSATVKTSSNQSPQNTANQ-----STTKTSNTYNDKSSSTIYSNETDK 145
 QY 159 PNLVEAYKA-----LKTTLBORAVNLEGLSSTAYNQIRNLDVLYNKASSLITKTLDPNG 214
 Db 146 SNLTOAKDVSTTPKTTIKPRLNRMAVNTVAAPQOQTNDVKHFS-----NI 194
 QY 215 GTLLDSNEI--TTANKNNITLSTINEQTN----- 243
 Db 195 DIAIDGHNQTTGKTFEATNSDVLKLNKNTIDDSVKEGDTFTFKYGGYFRPGRVRLP 254
 QY 244 -----ADALNSFIKKVIONNEQSFVGTFTNANVPSNY--SF-----V 280
 Db 255 SQONLYNAQGNLIAGIYDSTNTTYTFTNVDOQYTVRGSEFEQVAFKRNATDKT 314
 QY 281 AFSADVTPVNYKYARTVYN--GDEPSSRILANTNSITD-----VSWIYSLAGTNTKYQ 332
 Db 315 AYKMEVTLGNDTYSEIIVDYGKKAQPLISSNYINNEEDLSRNTAYVNPQNTYTKQT 374
 QY 333 FFSNYPGPGTYLYFP-----YKLVKAADANNVGLQYKLNNGVQOV--EFATSTSNNTT 386
 Db 375 FVTN-----LTGYKFNPAKNFKIYEVDNQFVDSFTPDTSKLKDVTDQFDVIYSNDNKT 430
 QY 387 ANTPAVDEIK 397
 Db 431 A-----TVDLWK 437

RESULT 11
 US-10-647-057-1
 ; Sequence 1, Application US/10647057
 ; GENERAL INFORMATION:
 ; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
 ; APPLICANT: NAGARAJA, R.
 ; APPLICANT: STEWART, GEORGE
 ; APPLICANT: NARAYANAN, SANJEEV
 ; APPLICANT: CHENGAPPA, M.
 ; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: PCT-30962
 ; CURRENT APPLICATION NUMBER: US/10/647,057
 ; CURRENT FILING DATE: 2003-08-22
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3241
 ; TYPE: PRT
 ; ORGANISM: Fusobacterium necrophorum
 US-10-647-057-1

Query Match 5.3%; Score 123; DB 6; Length 3241;
 Best Local Similarity 22.0%; Pred. No. 40;
 Matches 85; Conservative 61; Mismatches 161; Indels 80; Gaps 16;
 QY 63 GCMSTFKKDANPNNGOTOLEAARMELTDLINAKAMTLASLDQYAKTEASL-----SSAYSE 118
 Db 1431 GAVSVNKID--NDVEASVDKSSIEGANEIVAKDVKGSSDLAKYQALLNGKKKYLE 1487
 QY 119 AETVNNNLNA--TLEQLKMAKTN-----LESAINQANTDKTTFD-----NEHPNLVEA-- 164
 Db 1488 DRGINTGNGYVTKKEQLEKAKKKEGAVIVNAALSVAQDKSAGGVAIAVNTVKNKFKEL 1547
 QY 165 ----YKALKTLEQATNLEGLSSTAYNQIRNLDVLYNKASSLITKTLDPNGGTL--L 218
 Db 1548 SCSNKEAGEDKTHAKHVNYEAKSSVTVVNAASGLAISKDAFSCMGSGAWQDLSNDTIKAV 1607
 QY 219 DSNEIT--TANKNNITLSTINEQTNADALN-----SPIKKVIONNEOS-----F 263
 Db 1608 DKGRISADSLAVNANNISILGVNVAGTIAGSLSTAVGAAPANTTLHNKYSALITGKVNP 1667
 QY 264 VGTFTNANVQPSNYSFVAFSADVTVPVNYKA----RRTVWNGDEPSSRILANTN----- 313
 Db 1668 SGKNTKVNYQALNDSHITNVSGAGGAAIKQAGIGGMVSVNRGSDTEALVSDSEFEGVSS 1727

QY 314 ---SITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADAN--NVGLQ----- 363
 Db 1728 FNVDAKDQKTINTIAGN-----ANGGKAAG-----VCATVAHTNIGKQSVIAIV 1771
 QY 364 --YKLNNGNVQVEFATSTSNANTTAN 388
 Db 1772 KNSKITANDQDRKNINVTAKDYTMN 1798
 RESULT 12
 PCT-US02-37235-44
 ; Sequence 44, Application PC/TUS0237235
 ; GENERAL INFORMATION:
 ; APPLICANT: Handfield, Martin
 ; APPLICANT: Hillman, Jeffrey
 ; APPLICANT: Progulskie-Pox, Ann
 ; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigen
 ; TITLE OF INVENTION: in the Diagnosis, Treatment, and Monitoring of Periodontal Di
 ; FILE REFERENCE: MBH01-662A
 ; CURRENT APPLICATION NUMBER: PCT/US02/37235
 ; CURRENT FILING DATE: 2003-06-16
 ; NUMBER OF SEQ ID NOS: 234
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 1531
 ; TYPE: PRT
 ; ORGANISM: Actinobacillus actinomycetemcomitans
 PCT-US02-37235-44
 Query Match 5.2%; Score 122; DB 1; Length 1531;
 Best Local Similarity 21.8%; Pred. No. 24;
 Matches 102; Conservative 60; Mismatches 148; Indels 158; Gaps 24;
 QY 15 ILYGTNSPSTONV-----TSREVVSSVOLSEESTF-YLCPPPVGVSVIRLEFGCMS 66
 Db 329 ISLGANAKASQDNVVAIGYATATSGSMAIGOGAKSTFNSIALGTGTIVNSVDGGQSK 388
 QY 67 ITKDDANPNNGOTOLEAARMELTDLINAKAMTLASLDQYAKTEASLSEAYSEATVNNNL 126
 Db 389 FTAQNYDANNVAVANAGKE--RRIN-----VAGGRNDTDVYN--- 426
 QY 127 NATLEQLKMAKTNLESAL-----NQANTDKTT-----FNEHPNLVEAYKALKTLEQRA 176
 Db 427 ---IAQLKEFVNDNLAKSIAGAGYNGYETDGHYKAPVFSIKNTN---YHDVKTAVEAAQ 479
 QY 177 TNLGLSSTAYNQIRNLDVLYNKASSLITKTLDPNGGTLTLDSEITANKNNITLST 236
 Db 480 TNYVSVNST-----NTAADSNDYDKGAKAVGSI--A 508
 QY 237 INEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARR 296
 Db 509 LGEKATTTGTAAMNS-----IAIGLSNV--SQONTVALGANITATT----- 547
 QY 297 TVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPST--GYLY--FPYKLV 352
 Db 548 ---NG---SVILGNSST-----TEGSHPVSVNSATVNGTYTSGFTGVK 586
 QY 353 KAADANNVGLQYKLNNGNVQVEFATSTSNANTTANPTPAVDEIKVAKI-----VL 403
 Db 587 ESGHFVSI-----SKGNERQIK---NVAAGNVANSTNDVANGSQLFAVSRVEQGWIT 638
 QY 404 SGLRFG--QN-----TIELSVPTGEGNMKNKVPAMIG--NIYLSNNEN 442
 Db 639 SCVENGGTQNGAASATATIKPS-----NQVKLLAGKNLAVKQNCN 678
 RESULT 13
 PCT-US03-20460-14
 ; Sequence 14, Application PC/TUS0320460
 ; GENERAL INFORMATION:
 ; APPLICANT: Iowa State University Research Foundation, Inc. and Department of
 ; APPLICANT: Agriculture for

; PRIOR APPLICATION NUMBER: 60/098,427

S-10-472-078

; PRIOR APPLICATION NUMBER: 60/098,427

US-10-472-078-8

```

Query Match      5.1%  Score 119.5;  DB 6;  Length 1386;
Best Local Similarity 21.0%;  Pred. No. 29;
Matches 94;  Conservative 72;  Mismatches 154;  Indels 127;  Gaps 23;

QY 35 SSVQLSEESTFYLCPPPGVSTVIRLEFGCM---SITKKDANPNNGQTQLEAAARMELTDL 91
Db 733 SNVSI TEVE-----GTDTKLEYADAGNITVEGNGVTILGSNK---AKTHITKN 778

QY 92 INAKA---MTLAS-----LDYAKIEASLSS-----AYSEAET 121
Db 779 VSVKSNANLTSSANFNHVHGALTIGGSANIQGNLTANGDTVEAGDIVVSDDAKFAET 838

QY 122 VNN-NLNAATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-VEAYKALKTTLEQRATNL 179
Db 839 KNNLNIITGTNNGTSEINIKQGVVKLOGDIT--NNGNLNITTNASVNOKTIINGNITNK 896

QY 180 EG-----LSSTAYNQIRNNLVLYNKASSLITKTLPLNGGTLLD-----SNEITTA 226
Db 897 KGD LNIKDIKANAELIQIGNI-----SQKEGNLTISSDKINITKRIEKADTDQGNDSGV 952

QY 227 NKNINNTLSTINEOKTNADALSNSFIKK---VIONNEQSFVGTFTNANVQPSNYSFVAFS 283
Db 953 ASNANLTIKT--KELTLDNLNISGFENKAEITAKDSDLIIGKASSDN---SNAKQITFD 1007

QY 284 ADVTPVNYKYARRTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFNYGPGSTG 343
Db 1008 -----KVKDSKISAGNHNT-----LNSKVE--TSNSDGGSTG 1037

QY 344 YLYFPYKLVKAADANNVGLQYKLN---NGNV---QOVEFATS-----TSANNTTANPTPA 392
Db 1038 -----NGSDDNIGLTISAKDVTVNSNITSHKTVNISASEGGITTKAGTTINATTG 1088

QY 393 VDEIKVAKIVLSGLRFGONTIELSVPT 419
Db 1089 SVEVTAKTGDISGTISGK-TVSVTATT 1114

```

Search completed: October 8, 2003, 17:25:44
Job time : 11.6113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:03:00 ; Search time 182.64 Seconds
(without alignments)
2271.804 Million cell updates/sec

Title: US-09-147-052-2
Perfect score: 2324
Sequence: 1 MHYFRNCIFFLIVLYGYN.....SSNENADKIPGRRPGTEL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA_Main:*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2324	100.0	456	15 US-09-147-052-2 Sequence 2, Appl1

2	2317	99.7	456	24	US-09-901-572A-3	Sequence 3, Appl1
3	2015	86.7	1086	15	US-09-147-052-4	Sequence 4, Appl1
4	1929.5	83.0	384	27	US-10-131-591A-12	Sequence 12, Appl1
5	1777	76.5	357	27	US-10-131-591A-23	Sequence 23, Appl1
6	750.5	32.3	645	22	US-09-791-537-105581	Sequence 105581,
7	749	32.2	650	22	US-09-791-537-116408	Sequence 116408,
8	715.5	30.8	647	4	US-08-014-631-7	Sequence 7, Appl1
9	715.5	30.8	647	6	US-08-230-312-7	Sequence 7, Appl1
10	715.5	30.8	647	22	US-09-791-537-7385	Sequence 7385, Ap
11	715.5	30.8	647	22	US-09-791-537-34555	Sequence 34555, A
12	320	13.8	62	27	US-10-131-591A-5	Sequence 5, Appl1
13	314	13.5	885	3	US-07-722-860-13	Sequence 13, Appl1
14	314	13.5	885	6	US-08-213-449A-13	Sequence 13, Appl1
15	314	13.5	885	6	US-08-213-449B-13	Sequence 13, Appl1
16	308	13.3	62	27	US-10-131-591A-6	Sequence 6, Appl1
17	187.5	8.1	1095	21	US-09-710-279-3154	Sequence 3154, Ap
18	185.5	8.0	6641	28	US-10-282-122A-70580	Sequence 70580, A
19	185.5	8.0	10182	26	US-10-092-411A-3159	Sequence 3159, Ap
20	185.5	8.0	10203	18	US-09-450-969-4098	Sequence 4098, Ap
21	177	7.6	1947	12	US-08-827-356-5566	Sequence 5566, Ap
22	177	7.6	1947	20	US-09-611-529-4400	Sequence 4400, Ap
23	177	7.6	1947	25	US-09-950-084-4400	Sequence 4400, Ap
24	177	7.6	6713	28	US-10-282-122A-43811	Sequence 43811, A
25	176.5	7.6	1291	32	US-60-257-931-3393	Sequence 3393, Ap
26	176.5	7.6	1291	32	US-60-269-308-4414	Sequence 4414, Ap
27	176.5	7.6	6281	28	PCT-US02-03987-12996	Sequence 12996, A
28	176.5	7.6	6281	1	PCT-US02-03987-12996	Sequence 12996, A
29	176.5	7.6	6281	23	US-09-815-242-12996	Sequence 12996, A
30	176.5	7.6	6281	26	US-10-072-851-12996	Sequence 12996, A
31	175	7.5	3351	12	US-08-827-356-5564	Sequence 5564, A
32	175	7.5	3351	20	US-09-611-529-4562	Sequence 4562, Ap
33	175	7.5	3351	25	US-09-950-084-4562	Sequence 4562, Ap
34	174	7.5	2086	1	PCT-US02-03987-5639	Sequence 5639, Ap
35	174	7.5	2086	23	US-09-815-242-5639	Sequence 5639, Ap
36	174	7.5	2086	26	US-10-072-851-5639	Sequence 5639, Ap
37	174	7.5	2086	32	US-60-242-578-948	Sequence 948, App
38	174	7.5	2086	32	US-60-253-625-2292	Sequence 2292, Ap
39	174	7.5	2086	32	US-60-257-931-3208	Sequence 3208, Ap
40	174	7.5	2086	32	US-60-269-308-4231	Sequence 4231, Ap
41	174	7.5	5795	1	PCT-US02-03987-12610	Sequence 12610, A
42	174	7.5	5795	23	US-09-815-242-12610	Sequence 12610, A
43	174	7.5	5795	26	US-10-072-851-12610	Sequence 12610, A
44	173.5	7.5	936	29	US-10-311-879-29	Sequence 29, Appl1
45	173.5	7.5	933	14	US-09-095-272-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-147-052-2
; Sequence 2, Application US/09147052
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, Noboru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; PRIOR FILING DATE: 1999-04-05
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Query Match 100.0%; Score 2324; DB 15; Length 456;
 Best Local Similarity 100.0%; Pred. No. 3.2e-189;
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVTSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
 DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTVTSREVSSVOLSEESTFYLCPPPGVSTVIRL 60

QY 61 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
 DB 61 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120

QY 121 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
 DB 121 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180

QY 181 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 240
 DB 181 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 240

QY 241 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 300
 DB 241 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 300

QY 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 360
 DB 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 360

QY 361 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 DB 361 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420

QY 421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456
 DB 421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456

RESULT 2
 US-09-901-572A-3
 ; Sequence 3, Application US/09901572A
 ; GENERAL INFORMATION:
 ; APPLICANT: Nippon Zeon Co., Ltd.,
 ; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
 ; FILE REFERENCE: J209
 ; CURRENT APPLICATION NUMBER: US/09/901,572A
 ; CURRENT FILING DATE: 2003-03-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: Mycoplasma gallisepticum
 ; FEATURE:
 ; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
 US-09-901-572A-3

Query Match 99.7%; Score 2317; DB 24; Length 456;
 Best Local Similarity 99.6%; Pred. No. 1.3e-188;
 Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTVTSREVSSVOLSEESTFYLCPPPGVSTVIRL 60
 DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTVTSREVSSVOLSEESTFYLCPPPGVSTVIRL 60

QY 61 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
 DB 61 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120

QY 121 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
 DB 121 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180

QY 181 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 240

DB 181 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 240

QY 241 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 300
 DB 241 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 300

QY 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 360
 DB 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 360

QY 361 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 DB 361 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420

QY 421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456
 DB 421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456

RESULT 3
 US-09-147-052-4
 ; Sequence 4, Application US/09147052
 ; GENERAL INFORMATION:
 ; APPLICANT: SAITOH, Shuji
 ; APPLICANT: TSUZAKI, Yoshinari
 ; APPLICANT: YANAGIDA, Noboru
 ; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
 ; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
 ; FILE REFERENCE: 981167
 ; CURRENT APPLICATION NUMBER: US/09/147,052
 ; CURRENT FILING DATE: 1999-04-05
 ; PRIOR APPLICATION NUMBER: JP 08-103548
 ; PRIOR FILING DATE: 1996-03-29
 ; PRIOR APPLICATION NUMBER: PCT/JP97/01084
 ; PRIOR FILING DATE: 1997-03-28
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1086
 ; TYPE: PRT
 ; ORGANISM: hybrid
 US-09-147-052-4

Query Match 86.7%; Score 2015; DB 15; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 3.2e-162;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
 DB 691 EFGCMSITKKDANPNNGQTLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 750

QY 121 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
 DB 751 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 810

QY 181 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 240
 DB 811 GLSSTAYNQIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSTINEQ 870

QY 241 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 300
 DB 871 KTNADALSNSFTKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARTVWN 930

QY 301 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 360
 DB 931 GDEPSSRILANTNSITDVSIIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNV 990

QY 361 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 DB 991 GLOYKLNGNVOQVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050

QY 421 EGNMKNVAPMIGNIYLSNENNADKIPGYRRPGTFL 456

```

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79.
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified Ttl-1 portion (downstream of Bali) of

```

Query Match	32.3%	Score 750.5;	DB 22;	Length 645;
Best Local Similarity	40.6%;	Pred. No. 1.7e-54;		
Matches 177; Conservative	67;	Mismatches 133;	Indels 59;	Gaps 12;
QY	53	VGSTVIRLEFGCWSITKK-----DANPNNGO-----TGLEAARMELTDLINA	94	
Db	15	IGSFVMLAAASCCTTPNPNNPPSGMGGDTNFCGCGCMNRAASQELAAARMGLTTVFDS	74	
QY	95	KAMTLASLOQYAKIEIASLSAYSEATPVNNLNATLEQLKWKNTLESAINQANTDKTTF	154	
Db	75	KAKNLGLIVDYKKTQDTLTTRAYDAAKTVLDNSSTTQNLEAKRTRLETARTATSQTFS	134	
QY	155	DNEHPNVVEAYKALKTKTEQRATNLEGLSTAVNOIRNNLVLYNKASSLIITTKTDLPLNG	214	
Db	135	DEQHAELVKYYEELKITLSNETATLPYAAQYAGIKHLSGLGYDCAKITTTTLEPVEG	194	
QY	215	GTLLDSNEIHTANKNIINTLI--STINPQKTNADAALSNSFKVIQNNQESFVGFTTMANY	272	
Db	195	DP-LTADVMNMATKIYEAIKDEVILNPQKNATKLADSFKVLVKEKITGVBEAHN-KA	252	
QY	273	OPSNYSKVATSDVTVP-----NYKYARRTW-NGDEPSSRIILANT-----NSI	315	

Db 414 --EKVSPMIGNMYLSSSPNNWNI 435

RESULT 9
US-08-230-312-7
; Sequence 7, Application US/08230312
; GENERAL INFORMATION:
; APPLICANT: Browning, Glenn F.
; TITLE OF INVENTION: LIVE VACCINE VECTORS BASED ON MYCOPLASMA
; TITLE OF INVENTION: GALLISEPTICUM AND NUCLEIC ACID PROBES THEREFROM
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,312
; FILING DATE: April 20, 1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8801Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-230-312-7

Query Match 30.8%; Score 715.5; DB 6; Length 647;
Best Local Similarity 39.2%; Pred. No. 1.7e-51;
Matches 174; Conservative 63; Mismatches 134; Indels 73; Gaps 13;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90
DB 15 IGSFVMLAAASCTTPTNPPTNPSPGGMGGDTPGQGMNNAASQELAAARMGLTT 74

QY 91 LINAKAMTLASLDYAKIEASLSAYSEAEYNNLNATLEQKMAKTNLESAINOATD 150
DB 75 IFDSKAKNLGLVDYKKTQNTLTAYDAAKTVLDNSSSTQNLNEAKTRLEAIRTAATS 134

QY 151 KITTFDNEHNLVEAYKALKTTLEQKATNLEGLSSSTAYNOIRNLDVLYNKASSLTKTLD 210
DB 135 KOTFDEQHAELVKVYKELKTTLSNETATLAPYADAQYAGIKMHLGLYDAGKAITTKLE 194

QY 211 PLNGGTLLDSNEITTANKNINNTL--STINEQKN-----ADALSNSFIKKVIONNEQSFV 264
DB 195 PVGGDP-LTASAVMANTKIVEAIKDEVLPKONATKLADSLSSIVKKTITGVVEE---- 249

QY 265 GFTTNANQPSNYSFVAFSADVTVP-----NRYKARVTW-NGDEPSSRLANT--- 312
DB 250 ---AHNKAQPNYSFVGKRWYTELLELDKQVFPNWDYAORTIFTSNDEP--RSISNTPAD 304

QY 313 -----NSITDVSWIYSLAGTNTKYQSFSGNYGPGSTGYLPYKLVKAADANNVGLQYKLN 367
DB 305 GQTMAQPLSNVSWIYSLAGTNGAKYKLTLEFYYGPGSTGYLPYKLVNTSDQVKGLEYKLN 364

QY 368 NGNVOQVEPATSTSA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
DB 135 KOTFDEQHAELVKVYKELKTTLSNETATLAPYADAQYAGIKMHLGLYDAGKAITTKLE 194

QY 211 PLNGGTLLDSNEITTANKNINNTL--STINEQKN-----ADALSNSFIKKVIONNEQSFV 264
DB 195 PVGGDP-LTASAVMANTKIVEAIKDEVLPKONATKLADSLSSIVKKTITGVVEE---- 249

QY 265 GFTTNANQPSNYSFVAFSADVTVP-----NRYKARVTW-NGDEPSSRLANT--- 312
DB 250 ---AHNKAQPNYSFVGKRWYTELLELDKQVFPNWDYAORTIFTSNDEP--RSISNTPAD 304

QY 313 -----NSITDVSWIYSLAGTNTKYQSFSGNYGPGSTGYLPYKLVKAADANNVGLQYKLN 367
DB 305 GQTMAQPLSNVSWIYSLAGTNGAKYKLTLEFYYGPGSTGYLPYKLVNTSDQVKGLEYKLN 364

QY 368 NGNVOQVEPATSTSA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
DB 135 KOTFDEQHAELVKVYKELKTTLSNETATLAPYADAQYAGIKMHLGLYDAGKAITTKLE 194

Db 365 D-----ATKPSAITFGSDQTMNGKTPTVNDINVAKYTLANLNFSGNKIEFSVPA---- 413

QY 423 NMNKVAPMIGNIYLSNENNADKI 446
DB 414 --EKVSPMIGNMYLSSSPNNWNI 435

RESULT 10
US-09-791-537-7385
; Sequence 7385, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7385
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
US-09-791-537-7385

Query Match 30.8%; Score 715.5; DB 22; Length 647;
Best Local Similarity 39.2%; Pred. No. 1.7e-51;
Matches 174; Conservative 63; Mismatches 134; Indels 73; Gaps 13;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90
DB 15 IGSFVMLAAASCTTPTNPPTNPSPGGMGGDTPGQGMNNAASQELAAARMGLTT 74

QY 91 LINAKAMTLASLDYAKIEASLSAYSEAEYNNLNATLEQKMAKTNLESAINOATD 150
DB 75 IFDSKAKNLGLVDYKKTQNTLTAYDAAKTVLDNSSSTQNLNEAKTRLEAIRTAATS 134

QY 151 KITTFDNEHNLVEAYKALKTTLEQKATNLEGLSSSTAYNOIRNLDVLYNKASSLTKTLD 210
DB 135 KOTFDEQHAELVKVYKELKTTLSNETATLAPYADAQYAGIKMHLGLYDAGKAITTKLE 194

QY 211 PLNGGTLLDSNEITTANKNINNTL--STINEQKN-----ADALSNSFIKKVIONNEQSFV 264
DB 195 PVGGDP-LTASAVMANTKIVEAIKDEVLPKONATKLADSLSSIVKKTITGVVEE---- 249

QY 265 GFTTNANQPSNYSFVAFSADVTVP-----NRYKARVTW-NGDEPSSRLANT--- 312
DB 250 ---AHNKAQPNYSFVGKRWYTELLELDKQVFPNWDYAORTIFTSNDEP--RSISNTPAD 304

QY 313 -----NSITDVSWIYSLAGTNTKYQSFSGNYGPGSTGYLPYKLVKAADANNVGLQYKLN 367
DB 305 GQTMAQPLSNVSWIYSLAGTNGAKYKLTLEFYYGPGSTGYLPYKLVNTSDQVKGLEYKLN 364

QY 368 NGNVOQVEPATSTSA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEG 422
DB 365 D-----ATKPSAITFGSDQTMNGKTPTVNDINVAKYTLANLNFSGNKIEFSVPA---- 413

QY 423 NMNKVAPMIGNIYLSNENNADKI 446
DB 414 --EKVSPMIGNMYLSSSPNNWNI 435

RESULT 11
US-09-791-537-34555
; Sequence 34555, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF

Db 1 MHVFRNCIFFLIVILYGTNSSPSTQNTVSREVSSVOLSEESTFYLCPPPVGVSTVIRL 60

QY 61 EF 62
11

Db 61 EF 62

RESULT 13
US-07-722-860-13
: Sequence 13. Application US/07722860

GENERAL INFORMATION:
APPLICANT: NAZERIAN, Keyvan
APPLICANT: LEE, Lucy F.
APPLICANT: YANAGIDA, Noboru
APPLICANT: OGAWA, Ryohiei
APPLICANT: LI, Yi
TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
PROTECTION AGAINST MAREK'S DISEASE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 301 North Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747

```

;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
US-07-722-860-13

Query Match      13.5%; Score 314; DB 3; Length 865;
Best Local Similarity 100.0%; Pred. No. 5,6e-17;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1 MHYFRNCJFFLVILVLYGTNSPSTQNTVSREWSVQLSEESTFYLCPPPVGSVTIRL 60
        |||||
Db      1 MHYFRNCJFFLVILVLYGTNSPSTQNTVSREWSVQLSEESTFYLCPPPVGSVTIRL 60
        |||||

```

```

: Sequence 13, Application US/08213449A
:
: GENERAL INFORMATION:
:
: APPLICANT: NAZERIAN, Keyvan
:
: APPLICANT: LEE, Lucy F.
:
: APPLICANT: YANAGIDA, Noboru
:
: APPLICANT: OGAWA, Ryohei
:
: APPLICANT: LI, Yi
:

```

```

: TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
:
: TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
:
: NUMBER OF SEQUENCES: 19
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSER: BIRCH, STEWART, KOLASCH & BIRCH
:
: STREET: P.O. Box 747
:
: CITY: Falls Church
:
: STATE: Virginia
:
: COUNTRY: USA
:
: ZIP: 22040-0747
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/213,449A
:
: FILING DATE: 15-MAR-1994
:
: CLASSIFICATION: 424
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Murphy Jr., Gerald M.
:
: REGISTRATION NUMBER: 28,977
:
: REFERENCE/DOCKET NUMBER: 1644-108P
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (703) 205-8000
:
: TELEFAX: (703) 205-8050
:
: INFORMATION FOR SEQ ID NO: 13:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 865 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-213-449A-13

```

```
Query Match      13.5%; Score 314; DB 6; Length 865;  
Best Local Similarity 100.0%; Pred. No. 5.6e-17;  
Matches        61; Conservative    0; Mismatches   0; Indels     0; Gaps       0;
```

QY 1 MHYFRNCIFFLIVILVIGTNSPSTQNVTTSREVVSQQLESEESTFYLCPVPVGSTVIRL 60
| | | | | | | | | | | | | | | | | | | | | |
DB 1 MHYFRNCIFFLIVILVIGTNSPSTQNVTTSREVVSQQLESEESTFYLCPVPVGSTVIRL 60

QY 61 E 61
|
DB 61 E 61

```

RESULT 15
US-08-213-449B-13
; Sequence 13, Application US/08213449B
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, Noboru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213.449B

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 17:07:16 ; Search time 26.2345 Seconds
(without alignments)
2800.682 Million cell updates/sec

Title: US-09-147-052-2

Perfect score: 2324

Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNNENADKIPGRRPCTFL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	2324	100.0	456	9	US-09-147-052-2	Sequence 2, Appli
2	2317	99.7	456	12	US-09-901-572A-3	Sequence 3, Appli
3	2015	86.7	1086	9	US-09-147-052-4	Sequence 4, Appli
4	1929.5	83.0	384	15	US-10-131-591A-12	Sequence 12, Appli
5	1777	76.5	357	15	US-10-131-591A-23	Sequence 23, Appli
6	320	13.8	62	15	US-10-131-591A-5	Sequence 5, Appli
7	308	13.3	62	15	US-10-131-591A-6	Sequence 6, Appli
8	176.5	7.6	6281	9	US-09-815-242-12996	Sequence 12996, A
9	174	7.5	2086	9	US-09-815-242-5639	Sequence 5639, A
10	174	7.5	5795	9	US-09-815-242-12610	Sequence 12610, A
11	173.5	7.5	496	12	US-10-311-879-29	Sequence 29, Appli
12	173.5	7.5	936	12	US-08-781-986A-5249	Sequence 5249, Ap
13	170	7.3	520	15	US-10-056-052-2	Sequence 2, Appli
14	166.5	7.2	2659	12	US-10-311-879-28	Sequence 28, Appli
15	162	7.0	1029	9	US-09-815-242-5885	Sequence 5885, Ap

15	162	7.0	1048	9	US-09-815-242-13083	Sequence 13083, A
17	158	6.8	2434	9	US-09-815-242-5835	Sequence 5835, Ap
18	156	6.7	807	11	US-09-820-843A-108	Sequence 108, App
19	155.5	6.7	2478	9	US-09-815-242-5816	Sequence 5816, Ap
20	155.5	6.7	2478	9	US-09-815-242-12967	Sequence 12967, A
21	149.5	6.4	837	9	US-09-815-242-5883	Sequence 5883, Ap
22	149.5	6.4	875	9	US-09-815-242-13080	Sequence 13080, A
23	149.5	6.4	3158	9	US-09-815-242-12611	Sequence 12611, A
24	148.5	6.4	857	12	US-10-032-585-7857	Sequence 7857, Ap
25	145.5	6.3	2437	9	US-09-815-242-5834	Sequence 5834, Ap
26	145	6.2	1215	9	US-09-815-242-5908	Sequence 5908, Ap
27	145	6.2	1269	9	US-09-815-242-13113	Sequence 13113, A
28	144.5	6.2	596	11	US-09-946-37A-243	Sequence 243, App
29	144.5	6.2	596	12	US-10-015-387A-243	Sequence 243, App
30	144.5	6.2	596	12	US-10-063-735-100	Sequence 100, App
31	144.5	6.2	596	12	US-10-006-130A-243	Sequence 243, App
32	144.5	6.2	596	12	US-10-199-672-310	Sequence 310, App
33	144.5	6.2	596	12	US-10-006-172A-243	Sequence 243, App
34	144.5	6.2	596	12	US-10-187-749-310	Sequence 310, App
35	144.5	6.2	596	12	US-10-194-457-310	Sequence 310, App
36	144.5	6.2	596	12	US-10-184-642-310	Sequence 310, App
37	144.5	6.2	596	12	US-10-196-747-310	Sequence 310, App
38	144.5	6.2	596	12	US-10-015-392A-243	Sequence 243, App
39	144.5	6.2	596	12	US-10-017-253A-243	Sequence 243, App
40	144.5	6.2	596	12	US-10-173-689-310	Sequence 310, App
41	144.5	6.2	596	12	US-10-173-690-310	Sequence 310, App
42	144.5	6.2	596	12	US-10-173-691-310	Sequence 310, App
43	144.5	6.2	596	12	US-10-173-692-310	Sequence 310, App
44	144.5	6.2	596	12	US-10-173-694-310	Sequence 310, App
45	144.5	6.2	596	12	US-10-173-698-310	Sequence 310, App

ALIGNMENTS

RESULT 1

US-09-147-052-2
; Patent No. Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-2

Query Match 100.0%; Score 2324; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 9.5e-175;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHYFRNCIFFLIVLYGTNSSPSTQNTVSREVSVSVLSEESTFYLCPPVGVSVIRL	60
DB	1	MHYFRNCIFFLIVLYGTNSSPSTQNTVSREVSVSVLSEESTFYLCPPVGVSVIRL	60
QY	61	EFGCSITKDKANPNNGQTQLEAARMELDLINAKAMTLASLDYAKIEASLSAYSEAE	120
DB	61	EFGCSITKDKANPNNGQTQLEAARMELDLINAKAMTLASLDYAKIEASLSAYSEAE	120
QY	121	TVNNLNLTLEQLKMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEQRATNLE	180

Db 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
Qy 181 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 240
Db 181 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 240
Qy 241 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 300
Db 241 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 300
Qy 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 360
Db 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 360
Qy 361 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
Db 361 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
Qy 421 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456
Db 421 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456

RESULT 2

US-09-901-572A-3
; Sequence 3, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901.572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-09-901-572A-3

Query Match 99.7%; Score 2317; DB 12; Length 456;
Best Local Similarity 99.6%; Pred. No. 3.4e-174;
Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MHYFRRNCIFFLIVILYGTNSSPSTQNTSREVSVSSVOLSEESTFVLCPPPVGSTVIRL 60
Db 1 MHYFRRNCIFFLIVILYGTNSSPSTQNTSREVSVSSVOLSEESTFVLCPPPVGSTVIRL 60
Qy 61 EFGCMSITKKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
Db 61 EFGCMSITKKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
Qy 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
Db 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
Qy 181 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 240
Db 181 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 240
Qy 241 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 300
Db 241 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 300
Qy 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 360
Db 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 360
Qy 361 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420

Db 361 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
Qy 421 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456
Db 421 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456

RESULT 3

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1oru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147.052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCY/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

Query Match 86.7%; Score 2015; DB 9; Length 1086;
Best Local Similarity 100.0%; Pred. No. 7.1e-150;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 61 EFGCMSITKKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
Db 691 EFGCMSITKKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 750
Qy 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
Db 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 810
Qy 181 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 240
Db 811 GLSTAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLTLDSEITTTANKNNINNTLSTINEQ 870
Qy 241 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 300
Db 871 KTNADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWN 930
Qy 301 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 360
Db 931 GDEPSSRILANTNSITDVSWIYSLAGTNTKYQSFNSGPGSTGYLYFPYKLVKAADANNV 990
Qy 361 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
Db 991 GLOQKLNNGNVQVQEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
Qy 421 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 456
Db 1051 EGNMNVAPMIGNIYLSNNENNADKIPGYRRPGTFL 1086

RESULT 4

US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof

```

; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Query Match      83.0%; Score 1929.5; DB 15; Length 384;
Best Local Similarity 99.2%; Pred. No. 8.5e-144;
Matches 382; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 61 EFGCMSTTKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
DB 1 EF-CMSITTKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 59

QY 121 TVNNLNATLEOLKMAKTNLESAINQANTOKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
DB 60 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 119

QY 181 GLSSTAYNOIRNLVDLYNKASSLITKTLPLNGGTLSDNEITTANKNNINNTLSTINEQ 240
DB 120 GLSSTAYNOIRNLVDLYNKASSLITKTLPLNGGTLSDNEITTANKNNINNTLSTINEQ 179

QY 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKYARRVTWN 300
DB 180 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKYARRVTWN 239

QY 301 GDEPSSRILANTNSITDVSIIYSLAGNTKYQSFNSYGPSTGYLYPPYKLVKAADANNV 360
DB 240 GDEPSSRILANTNSITDVSIIYSLAGNTKYQSFNSYGPSTGYLYPPYKLVKAADANNV 299

QY 361 GLOKLNNGNVQVEFATSPANNTPANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
DB 300 GLOKLNNGNVQVEFATSPANNTPANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 359

QY 421 EGNMKNVAPMIGNIYSSNENNAK 445
DB 360 EGNMKNVAPMIGNIYSSNENNAK 384

RESULT 5
US-10-131-591A-23
; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 23
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of
; OTHER INFORMATION: pNZ40K-S
US-10-131-591A-23

Query Match      76.5%; Score 1777; DB 15; Length 357;
Best Local Similarity 98.6%; Pred. No. 7.7e-132;
Matches 352; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 89 TDLINAKAMTSLASLDYAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQAN 148
DB 89 TDLINAKAMTSLASLDYAKIEASLSAYSEATVNNLNATLEOLKMAKTNLESAINQAN 148

; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

Query Match      83.0%; Score 1929.5; DB 15; Length 384;
Best Local Similarity 99.2%; Pred. No. 8.5e-144;
Matches 382; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 61 EFGCMSTTKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 120
DB 1 EF-CMSITTKDANPNNGQTOLEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE 59

QY 121 TVNNLNATLEOLKMAKTNLESAINQANTOKTTFDNEHPNLVEAYKALKTTLEQRATNLE 180
DB 60 TVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLE 119

QY 181 GLSSTAYNOIRNLVDLYNKASSLITKTLPLNGGTLSDNEITTANKNNINNTLSTINEQ 240
DB 120 GLSSTAYNOIRNLVDLYNKASSLITKTLPLNGGTLSDNEITTANKNNINNTLSTINEQ 179

QY 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKYARRVTWN 300
DB 180 KTNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKYARRVTWN 239

QY 301 GDEPSSRILANTNSITDVSIIYSLAGNTKYQSFNSYGPSTGYLYPPYKLVKAADANNV 360
DB 240 GDEPSSRILANTNSITDVSIIYSLAGNTKYQSFNSYGPSTGYLYPPYKLVKAADANNV 299

QY 361 GLOKLNNGNVQVEFATSPANNTPANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
DB 300 GLOKLNNGNVQVEFATSPANNTPANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 359

QY 421 EGNMKNVAPMIGNIYSSNENNAK 445
DB 360 EGNMKNVAPMIGNIYSSNENNAK 384

RESULT 6
US-10-131-591A-5
; Sequence 5, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:
; OTHER INFORMATION: MDVgB signal
US-10-131-591A-5

Query Match      13.8%; Score 320; DB 15; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.1e-18;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVILYGTNSPSTQNTSVTSREVSVSSVOLSEESTFYLCPPPVGTVIRL 60
DB 1 MHYFRNCIFFLIVILYGTNSPSTQNTSVTSREVSVSSVOLSEESTFYLCPPPVGTVIRL 60

QY 61 EF 62
DB 61 EF 62

RESULT 7
US-10-131-591A-6
; Sequence 6, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131.591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Marek's disease gammaherpesvirus
; FEATURE:

```

```
; OTHER INFORMATION: Modified Vgb signal
US-10-131-591A-6

Query Match      13.3%; Score 308; DB 15; Length 62;
Best Local Similarity 96.8%; Pred. No. 2.7e-17;
Matches 60; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTQSPSTQNTVTSREVSVSVQSEESTFYLCPPPVGVSTVIRL 60
Db 1 MHYFRNCIFFLIVLYGTQSPSTQNTVTSREVSVSVQSEESTFYLCPPPVGVSTVIRL 60

QY 61 EF 62
Db 61 EF 62

RESULT 8
US-09-815-242-12996
; Sequence 12996, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match      7.6%; Score 176.5; DB 9; Length 6281;
Best Local Similarity 22.3%; Pred. No. 0.0005;
Matches 98; Conservative 66; Mismatches 178; Indels 97; Gaps 16;

QY 75 NNGOTLEAARMETDLINAKAMTLASLDQYAKIESLSAYSEAT-----VNNLNAT 129
Db 266 SNTSTQLTAMANLQNGINDKNTLAS-ENYHDADSDKKTAYTQAVTNAENILNKNSGSN 324
QY 130 LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQ 189
Db 325 LD-----KTAVENALSQVANKAGLNH-NLEQAKSNANTTI-----NGLQHLTAQKDK 374
QY 190 IRNRLVDLYNKAS-SLITKTLDPUNG--GTLDDSNELTANKNINNTLSTINEOKT----- 242
Db 375 LKQVQQAQNVAGVDYTKSSANTLNGANGTLRNSIQDNTATKNGQNYLDATERKNTYNN 434
QY 243 -----NADALSNSFIKKVIONNEQSFVGTFTNANVPNSYFVAFSADVT 287
```

```
Db 435 AVDSANGVINATSNPMDANAINQATQVTSTKNALDGTNHLTQAKQT-----ATNAIDG 489
QY 288 PNYKYARRTVWNGDESSRIILANTNSITDVSWIYSLA-----GNTKYOFSF 335
Db 490 ATNLAQKQKDALKAQVTSNQRVANVTSIQOTANELNTAMGLOHGDIDENATKQTKYRD 549
QY 336 SNYGSTGYLYFPYKLVRAADA-----NNVGLQYKLN-----NGNVQOYEF 376
Db 550 APOSKKTAY-----DQAVAAAKAILNKQTSNSDKAAVDRAALQQVTTSTKDALNGDAKLA 605
QY 377 ATTSAN-----NTTANPTPAVDIYKIVLSGLRFGQNTIELSVPTGNNM----- 425
Db 606 KAAKQNLGTLNHIITNAQRTDLEGINQATT-----DGVNTVKTNTNANTLDGAMNSLOGS 660
QY 426 ---KVAPMIGNIYLSNEN 441
Db 661 INDKDATLRNQNYLDADES 679

RESULT 9
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5639
; LENGTH: 2086
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5639

Query Match      7.5%; Score 174; DB 9; Length 2086;
Best Local Similarity 20.9%; Pred. No. 0.00016;
Matches 104; Conservative 77; Mismatches 194; Indels 122; Gaps 19;

QY 17 YGNTSSPSTQNTVTSREVSVSVQSEESTFYLCPPPVGVSTVIRLEFCMSTTKKDANPN 76
Db 886 HALNLTSSINNAQRDLTKI-----DQAT-----IVAGVEAVSNT 921
QY 77 GQTQLEAARMETDLINAKAMTLASLDQYAKIESLSAYSEAT-----VNNLNATLE 131
Db 922 G-TQLTAMANLQNGINDKNTLAS-ENYHDADSDKKTAYTQAVTNAENILNKNSGSLD 979
QY 132 QLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQIR 191
```



```
Db 980 -----KAAVENALSOVTKAGALNGNH-NLEQAKSNANTTI-----NGLOHLLTQAOKDKLK 1029
QY 192 NNLVDLYNKAS-SLITKTLDPNG--GTLLDSNEITTTANKNINNTLSTINEQKT-----242
Db 1030 QOVQQAQNVAGVDTVKSSANTLNGAMGTLRNSIQDNTATNNGQNYLDATESNKTNNNAV 1089
QY 243 -----NADALSNSFIKKVIONNEQSFVGTETNANVQPSNYSFVAFSADVTPV 289
Db 1090 DSANGVINATSNPNMDANAINOIAQTSTKALDGTNLTAQAKQT-----ATNAIDGAT 1144
QY 290 NYKYARRVTWNGDEPSSRILANTNSITDVSWIYSLA-----GTNTKYQFSFN 337
Db 1145 NLNKAQKDALKAQVTSQARVANVTISIQOTANELNTAMGLOHIGIDDENATKQTKYRDAE 1204
QY 338 YGPSTGYLYFPYKLYKAADA-----NNVGLQYKLN-----NGNVQVQEFAT 378
Db 1205 QSKKTAY----DQAVAAAKAILNKOTGNSDKAAVDRLAQVTSKDALNGDAKLAEAKA 1260
QY 379 STSAN-----NTTANPTPAVDKIVAKIVLSGLRFGQNTIELSVPTGEGNWN-----425
Db 1261 AAKQNLGLTNHITNAQRTALEGOINQATTV-----DGVNTVKTNTANTLDGAMNSLQGSIN 1315
QY 426 -KVAPMIGNIYLSNEN 441
Db 1316 DKDATLRNQNYLDADES 1332
```

RESULT 10

```
US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12610
; LENGTH: 5795
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12610
```

Query Match 7.5%; Score 174; DB 9; Length 5795;
Best Local Similarity 20.9%; Pred. No. 0.0007;
Matches 104; Conservative 77; Mismatches 194; Indels 122; Gaps 19;

QY 17 YGTNSPSTQNTSEVSVVQLSEESTFYLCPPPGVGVSTVIRLEFGCMSTTKKDANPNN 76

```
Db 3163 HALNNLTINNAAKXKDLTKI-----DQAT-----TVAGVEAVSWT 3198
QY 77 GQTOLEAARMELTDLINAKAMTASLQDYAKIEASLSAYSEAE-----VNNNINLNTAE 131
Db 3199 G-TQLNTAMANLQNGINKANTLAS-ENYHDADSKKTAYTAQVTAENILNKNSSNLD 3256
QY 132 QLKMAKTNLESAINQANTDKTTFDNEHPNLVEYALKTKTLBQRATNLEGLSSTAYNQIR 191
Db 3257 -----KAAVENALSOVTKAGALNGNH-NLEQAKSNANTTI-----NGLOHLLTQAOKDKLK 3306
QY 192 NNLVDLYNKAS-SLITKTLDPNG--GTLLDSNEITTTANKNINNTLSTINEQKT-----242
Db 3307 QOVQQAQNVAGVDTVKSSANTLNGAMGTLRNSIQDNTATNNGQNYLDATESNKTNNNAV 3366
QY 243 -----NADALSNSFIKKVIONNEQSFVGTETNANVQPSNYSFVAFSADVTPV 289
Db 3367 DSANGVINATSNPNMDANAINOIAQTSTKALDGTNLTAQAKQT-----ATNAIDGAT 3421
QY 290 NYKYARRVTWNGDEPSSRILANTNSITDVSWIYSLA-----GTNTKYQFSFN 337
Db 3422 NLNKAQKDALKAQVTSQARVANVTISIQOTANELNTAMGLOHIGIDDENATKQTKYRDAE 3481
QY 338 YGPSTGYLYFPYKLYKAADA-----NNVGLQYKLN-----NGNVQVQEFAT 378
Db 3482 QSKKTAY----DQAVAAAKAILNKOTGNSDKAAVDRLAQVTSKDALNGDAKLAEAKA 3537
QY 379 STSAN-----NTTANPTPAVDKIVAKIVLSGLRFGQNTIELSVPTGEGNWN-----425
Db 3538 AAKQNLGLTNHITNAQRTALEGOINQATTV-----DGVNTVKTNTANTLDGAMNSLQGSIN 3592
QY 426 -KVAPMIGNIYLSNEN 441
Db 3593 DKDATLRNQNYLDADES 3609
```

RESULT 11

```
US-10-311-879-29
; Sequence 29, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: toxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-311-879-29
```

Query Match 7.5%; Score 173.5; DB 12; Length 496;
Best Local Similarity 22.5%; Pred. No. 2.2e-05;
Matches 111; Conservative 66; Mismatches 185; Indels 131; Gaps 23;

```
QY 2 HYFRNCIFFLIVIL-----YGTNSS-----PSTQNTSEVSVVQLSEESTFYLCPPP 52
Db 9 HAIRKKSIGVASVLVGTGLIGLLSKKEADASENVTSQDSASNESKSNDSVSAAPKT 68
QY 53 VGSTVIRLEFGCMSTTKKDANPNNGQTL--EAAARMELTDLINAKAMTASLQDYAKIEA 110
Db 69 DDTNV-----SDTKTSSNTNNGETSVAPNPAQOETTQSSSTNATT-----108
QY 111 SLSSAYSEAEVTNNNLNATLEQLKVAKTNLESAINQANTDKTTFDNEHPNLVEYALKAT 170
Db 109 EETPYTGEATTTTNTQANTPATTSNTNAEELVNO-TSNETTFND--TNTVSSVNS---162
QY 171 TLEQATNLEGLSST-----AYNOIRNLVLDLYNK-----ASSLIT 206
Db 163 ---PQNSTNAENVSTQDTSTEATPSNESAPOSTDASNKDVVNVQAVNTSAPRMRAPSLAA 220
```



```
Db 53 NPAQQTQSSSTNATT-----BETPVYGEATTTTNOANTPATQSSNTNAE 100
Qy 142 SAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLSST-----AYNQIRN 192
Db 101 ELVNO--TSNETTND--TNTVSSVNS-----PQNSTNAENVSTQDSTSTEATPSNSESAP 152
Qy 193 NLVDLYNK-----ASSLITKTLPLNGTGILLDSNEITANKNINNTLSTI 237
Db 153 QSTDASNKNVYVQAVNTSAPRMRAFSLAAVAADAPAAAGTDI--TNLTNTVTVGIDSGTIV 211
Qy 238 NEQ-----KTN--ADALSNSFIK-----KVIONNESQFVGTFTNANVQPSNFFVAFSADVTP 288
Db 212 PHQAGYVKNLNGVSPVNSAVKGFDTKIVPKELNNGVTSTAKVPP-----257
Qy 289 VNYKIARTVWNGDEPSSRILANTNSITDVSWIYSLAG--TNTKYQFSFSNYGPGSTGYLYF 347
Db 258 -----IMAGDQ-----VLANGVIDSGNVITFTDYVNTKDDVRATILTMPA--YI-- 300
Qy 348 PYKLVRKAADANNVQLYKLNNGVQVQVEFATSTSNANTTANPTPAVDEIKVAKI-----401
Db 301 -----DPENV-----KKTGNV-----TLATGIGSTTANKTVLVDYKRYGFYFNLISIK 342
Qy 402 -VLSGLRFGQNTIELSVPTGEGNMKNVAPMI--GNIYLSNEN 441
Db 343 GTIDQIDKNTNTYRQTIVVTPSGDNVIAPIVLTGMLKPNWDSN 384

RESULT 14
US-10-311-879-28
; Sequence 28, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: toxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2659
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-311-879-28

Query Match 7.2%; Score 166.5; DB 12; Length 2659;
Best Local Similarity 19.7%; Pred. No. 0.00086;
Matches 100; Conservative 90; Mismatches 209; Indels 109; Gaps 18;

Qy 17 YGTNNSPST-----QNVTSREVSVSSVOLSEESTFYLCPPPVGSTV 57
Db 817 HVSNSNPSTLTAPAAHTVNTTEIVKDYGSNVTAIEINNAVQVANKRT-----ATI 866
Qy 58 -----IRLEFGCMSITKDKANPNQGT-----QLEARMELTDLINAKA 96
Db 867 KNGTAMPTNLAGGSTTTPVTVTVNDGSTEEVQESIFTKADKRELITAKNHLDDPVSTEG 926
Qy 97 MTLASLDQY-----AKIEASLSAYSEAETVNNNLTLEOLKMAKTNLESAINQANTDKT 152
Db 927 KPQFTIYNNAMHNAQQOINTATEAQOVINNERATPQQVSDALTKVRAAQTKIDQAKA 986
Qy 153 TFDNEHPN--LVEAYKALKTTLEQRATNLEGLSSTAYNQIRNNLVLDYN-----KASSLIT 206
Db 987 LLQNKEDNSQLVTSKNLQSSVNO-----VPSTA--GMTQOSIDNYNAKKREATEIT 1037
Qy 207 KTLPLNGTGILLDSNEITANKNINVLSTINEQKTNADALSNSFIKVIQNNQSFVGT 266
Db 1038 AAQRVIDNGD--ATAQQISDEKHRVDNALTNALNQAKHDLTADTHALEQAVQQLNR---TGT 1093
Qy 267 FTN---ANVQPSNTSFAVFSADVT-----PV-NYKYARRRTVWNGDEPSSRILA 310
Db 1094 TTGKKPASITAYNSIRALQSLDLSAKNSANAIQKPIRTVQEVQSALTNNVRNRLTQ 1153
```

```
Qy 311 NTNSITDVSWIYSLAGTNTK-----YQFSFSNYGSGTGYLPYKLVKRAADANNVGLQYKL 366
Db 1154 AINOLVPLADNSALKTKTKILDEINKSVTTDGMTQSSIOQAYENAKRAGOTESTNAQNV 1213
Qy 367 NNGNVQVEFATS-----TSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVP 418
Db 1214 NNGDARTDQQTAAEYTKVEEKYNSLQKAIAGLTPDLAPLQTAQTQL-----QNDID--QP 1265
Qy 419 TGEQNMKNVAPMIGNIYLSNENNADKI 446
Db 1266 TSTTGMTSASIAAFNEKLSAARTKIQEI 1293

RESULT 15
US-09-815-242-5885
; Sequence 5885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5885
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5885

Query Match 7.0%; Score 162; DB 9; Length 1029;
Best Local Similarity 23.8%; Pred. No. 0.0005;
Matches 88; Conservative 53; Mismatches 146; Indels 82; Gaps 19;

Qy 79 TOLEARMELTDLINAKAMTLASLDQYAKIEASLSAYSEA--ETVNNNLTLEOLKMAK 137
Db 73 TDLNTAMGNLQALNDEQTTLNS--QNYQDATPSKKTATYNAVQAAKOILINKSNGQK--TK 130
Qy 138 TNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQIRNNLVLD 197
Db 131 DQVTEAMNQVNSAKNNLDG-----TRLDDQAKQAKAQQLNNHTLTAAKTNLTNQ-----181
Qy 198 YNKASSLITKTLPLNGTGILLDSNEITANKN--INTLSTINEQKTNADA--LSNSFI--252
Db 182 -----INSGTTVAGVTVQVQSNANTLDQAMNTRLQSIANDKATKASDYVDA 227
Qy 253 ---KKVIONNESQFVGTFTNANVQPS--NYSVFAFSADVTVPVNYKVARFTVNGDE-----303
Db 228 NNDKQATYNNAAAEETIINANSPEMNPSTITOKAE--QVN---SSKTLANGDENLFAA 282
```

[illegible]

Search completed: October 8, 2003, 17:28:02
Job time : 28.2345 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:59:40 ; Search time 12.8169 Seconds
(without alignments)
1505.341 Million cell updates/sec

Title: US-09-147-052-2

Perfect score: 2324

Sequence: 1 MHYFRNCIFFLVILYGTN.....SSNENADKIPGYRRPCTFL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	82.2	615	2	US-08-525-742-10
2	1856	79.9	610	2	US-08-525-742-8
3	1638	70.5	368	2	US-08-525-742-4
4	1612	69.4	368	1	US-08-185-851A-4
5	807	34.7	661	2	US-08-525-742-2
6	314	13.5	865	1	US-07-803-633A-13
7	185.5	8.0	10182	4	US-09-134-001C-3159
8	173.5	7.5	933	3	US-08-293-728-2
9	173.5	7.5	933	3	US-09-421-868-2
10	170.5	7.3	1041	1	US-08-220-151-4
11	170.5	7.3	1041	1	US-08-413-118-4
12	170.5	7.3	1041	3	US-08-473-446-4
13	154.5	6.6	682	4	US-08-836-687B-37
14	146.5	6.3	1073	3	US-09-541-782-6
15	146.5	6.3	1073	4	US-09-723-820-6
16	144.5	6.2	930	4	US-09-134-001C-5314
17	140.5	6.0	1002	4	US-09-268-347-24
18	140	6.0	2137	4	US-09-134-001C-4463
19	139	6.0	1004	4	US-09-268-347-30
20	136	5.9	943	3	US-08-911-321-4
21	136	5.9	1098	3	US-08-923-992A-8
22	133.5	5.7	903	3	US-08-804-439A-22
23	133.5	5.7	903	3	US-08-720-229-22
24	133.5	5.7	904	3	US-08-632-537-1
25	133.5	5.7	904	5	PCT-US96-05316-1
26	133.5	5.7	904	6	5244792-4
27	133.5	5.7	1073	4	US-09-206-942-49

28 133.5 5.7 1079 4 US-09-206-942-47 Sequence 47, Appl
29 132.5 5.7 903 1 US-08-220-151-8 Sequence 8, Appl
30 132.5 5.7 903 1 US-08-413-118-8 Sequence 8, Appl
31 132.5 5.7 903 3 US-08-473-446-8 Sequence 8, Appl
32 131.5 5.7 1164 3 US-08-923-992A-2 Sequence 2, Appl
33 131 5.6 1104 3 US-08-923-992A-4 Sequence 4, Appl
34 129.5 5.6 1095 4 US-09-206-942-45 Sequence 45, Appl
35 129.5 5.6 1101 4 US-09-206-942-43 Sequence 43, Appl
36 129 5.6 1095 4 US-09-206-942-69 Sequence 69, Appl
37 129 5.6 1536 1 US-08-038-682-2 Sequence 2, Appl
38 129 5.6 1536 1 US-08-302-832-2 Sequence 2, Appl
39 129 5.6 1536 2 US-08-530-198-2 Sequence 2, Appl
40 129 5.6 1536 2 US-08-469-880-2 Sequence 2, Appl
41 129 5.6 1536 2 US-08-728-470-2 Sequence 2, Appl
42 129 5.6 1536 2 US-08-617-697-2 Sequence 2, Appl
43 129 5.6 1536 3 US-08-719-641-2 Sequence 2, Appl
44 129 5.6 1536 4 US-09-206-942-67 Sequence 67, Appl
45 129 5.6 1833 4 US-08-621-944A-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-525-742-10
; Sequence 10, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroo
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyochito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930

```

; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-10

Query Match      82.2%   Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%   Pred. No. 4.3e-125;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 64 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTTLASQDYAKIEASLSAYSEAEVTN 123
Db 27 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTTLASQDYAKIEASLSAYSEAEVTN 86
QY 124 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 183
Db 87 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLA 146
QY 184 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDSDNEITTANKNINNTLSTINEQKTN 243
Db 147 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDSDNEITTANKNINNTLSTINEQKTN 206
QY 244 ADALNSFKIKVQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 303
Db 207 ADALNSFKIKVQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 304 PSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 363
Db 267 PSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 326
QY 364 YKLNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 423
Db 327 YKLNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 424 MNKVAPMIGNIYLSNNENADKI 446
Db 387 MNKVAPMIGNIYLSNNENADKI 409

RESULT 2
US-08-525-742-8
; Sequence 8, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroko
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemitsu
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-742-8

Query Match      79.9%   Score 1856; DB 2; Length 610;
Best Local Similarity 95.5%   Pred. No. 2.4e-121;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 64 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTTLASQDYAKIEASLSAYSEAEVTN 123
Db 27 CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTTLASQDYAKIEASLSAYSEAEVTN 86
QY 124 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 183
Db 87 NNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 146
QY 184 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDSDNEITTANKNINNTLSTINEQKTN 243
Db 147 STAYNOIRNLVDLYNKASSLITKTLPLNGGTLTLDSDNEITTANKNINNTLSTINEQKTN 206
QY 244 ADALNSFKIKVQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 303
Db 207 ADALNSFKIKVQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 304 PSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 363
Db 267 PSSRILANTNSITDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 326
QY 364 YKLNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 423
Db 327 YKLNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 424 MNKVAPMIGNIYLSNNENADKI 445
Db 387 MNKVAPMIGNIYLSNNENADKI 408

RESULT 3
US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroko
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemitsu

```

```

; APPLICANT: Takahashi, Kiyochito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESS: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-742-4

Query Match 70.5%; Score 1638; DB 2; Length 368;
Best Local Similarity 95.5%; Pred. No. 1.8e-106;
Matches 322; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 64 CMSITKDPANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEVYN 123
DB 27 CMSITKDPANPNNGQTQLEAARMELTDLINAKARTLASLDYAKTEASLSAYSEAEVYN 86
QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 183
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 184 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTANKNINNTLSTINEQKTN 243
DB 147 STAYNQIRNNLDVLYNKASSLITKTLDPNGGMLDSNEITVNRNINNTLSTINEQKTN 206
QY 244 ADALNSFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPNTKYARRTVWNGDE 303
DB 207 ADALNSFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPNTKYARRTVWNGDE 266
QY 304 PSSRILANTNSTIDVSWIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 363
DB 267 PSSRILANTNSTIDVSWIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 326
QY 364 YKLNNGNVQOQVEFATSTANNTTANPTQQLMRKLLK 400
; : : :

; 327 YKLNNGNVQOQVEFATSTANNTTANPTQQLMRKLLK 363
;
; RESULT 4
; US-08-185-851A-4
; Sequence 4, Application US/08185851A
; Patent No. 5489430
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Fujisawa, Ayumi
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemitsu
; TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
; TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
; TITLE OF INVENTION: Well As Vaccines Utilizing the Same
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESS: Naughton
; STREET: 1725 K Street, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
; SOFTWARE: ASCII from Word Perfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,851A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Theresa M. Stevens-Smith
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: PO-8-A930918
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-887-0357
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-185-851A-4

Query Match 69.4%; Score 1612; DB 1; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.2e-104;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 64 CMSITKDPANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEVYN 123
DB 27 CMSITKDPANPNNGQTQLEAARMELTDLINAKARTLASLDYAKTEASLSAYSEAEVYN 86
QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 183
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 184 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLDSNEITTANKNINNTLSTINEQKTN 243
DB 147 STAYNQIRNNLDVLYNKASSLITKTLDPNGGMLDSNEITVNRNINNTLSTINEQKTN 206
QY 244 ADALNSFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPNTKYARRTVWNGDE 303
DB 207 ADALNSFIKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTVPNTKYARRTVWNGDE 266
QY 304 PSSRILANTNSTIDVSWIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 363
DB 267 PSSRILANTNSTIDVSWIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 326
```

QY 364 YKLANGNVOQVEFATSTANNTTANPTPAVDKIVAK 400
 Db 327 YKLANGNVOQVEFATSTANNTTANPTQQLRLKLLK 363

RESULT 5

US-08-525-742-2
 ; Sequence 2, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hiroo
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemasa
 ; APPLICANT: Takahashi, Kiyochoito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSEE: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742
 ; FILING DATE: 25-SEP-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-074139
 ; FILING DATE: 31-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-245625
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP94/00541
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLeland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 950811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-659-2930
 ; TELEFAX: 202-8870357
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

US-08-525-742-2

Query Match 34.7%; Score 807; DB 2; Length 661;
 Best Local Similarity 41.6%; Pred. No. 2.6e-48;
 Matches 188; Conservative 64; Mismatches 128; Indels 72; Gaps 13;

QY 51 PPVGSVIRLEFCMSITKKDANPNNGOTQLEARMELTD-----LINAKMTIASLQ 103

Db 35 PPSGG-----NNGGNTNPSCQGMMAAKELADAKAALTTLINGETANLASYE 84

QY 104 DYAKIEASLSSAYSEAEATYNNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVE 163

Db 85 DYAKIKSELTSAKETAKAVSAKTGATLNEYNEAKTTLDAAIKKAASAKNDFDAHGSLVE 144
 QY 164 AYKALKTTLQORATNLEGLSSTAYNQIRNLVDLYNKASSLITKTLPLNGTLLDSNEI 223
 Db 145 AYNLKEETKEETNLDLANENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPEVMSV 203
 QY 224 TTANKINNTLSTINEOKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFYAFS 283
 Db 204 TQANQDITNATSLRIANKQADNLANFSFIQSLVKNLNRVDVANNOE-QPANYSFVGF 262
 QY 284 ADTPVNYKYARTVWNGDE-----PSSRLANTNSITDVSWIYSLAGTNTKYQFSGN 337
 Db 263 VNVDTPNWFQAQRKVAASENTPLATTPAEDATQQAASLTQVSWIYSLAGTNTKYQFSGN 322
 QY 338 YG-PSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVPEF-----ATTSAN--NTTA 387
 Db 323 FGAEKYALYFPYKLVKATSD--NVGLQYKLNNGTQKQINFVQTPASGSSDVAANEETMA 380
 QY 388 NP-----TPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSS---- 438
 Db 381 SPAEMQSAPTVDIDIKIAKVALSNLKFNSNTIEFSVPTG-----KAAPMIGNIYLSSNSE 435
 QY 439 -----NENN-----ADKIPGY 449
 Db 436 VKNKNIYDDLFGNSFNENNPTATVTDLLRGY 467

RESULT 6

US-07-803-633A-13
 ; Sequence 13, Application US/07803633A
 ; Patent No. 5369025
 ; GENERAL INFORMATION:
 ; APPLICANT: NAZERIAN, Keyvan
 ; APPLICANT: LEE, Lucy F.
 ; APPLICANT: YANAGIDA, No. 5369025oru
 ; APPLICANT: OGAWA, Ryohei
 ; APPLICANT: LI, Yi
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
 ; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: 301 No. 5369025th Washington Street
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/803,633A
 ; FILING DATE: 19911210
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M.
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 1644-103P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 241-1300
 ; TELEFAX: (703) 241-2848
 ; TELEX: 248345
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 865 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-803-633A-13


```
Dd      4433 VPKTQLQEAKNLENSINOOTDGTQSLNNYNDKAKARQNLEKISKVL-----GG 4487

Qy      410 QNTI-ELSVPTGEGNMKV-----PMIGNIYLSSNENNAKD 445
           ||| : | | : | | : | | : | | : | | : | | :
Db      4488 OPTVAETIQWTDERANAHKQALDTARSOLTLNRREPIINHINESHLNNAOK 4537


RESULT 8
US-08-293-728--2
; Sequence 2, Application US/08293728D
; Patent No. 608341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; ATTORNEY: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728--2

Query Match          7.5%   Score 173.5; DB 3; Length 933;
Best Local Similarity 22.5%; Pred. No. 0.0005;
Matches 111; Conservative 66; Mismatches 185; Indels 131; Gaps 23;

Qy    2 HYFRNCIFFLIVL-----YGNNSS-----PSTQNTVSRVVVSQLSSEESTFYLCPPP 52
       ||| : | | : | | : | | : | | : | | : | | : | | :
Db     9 HAIRKKSIGVASLVGLTGIFGLLLSKKEADASENSVTQSDASNESKSNDSSVSAAAPT 68
```

```

QY 53 VGSFVHLEFGCMSITKKDANPNNGQOL--EARMELTDLINAKAMTASLQDYAKIEA 110
Db 69 DDENV-----SDTKSSNTNGETSVAQNPAQQUETTOSSSTNATT----- 108
QY 111 SLSAYSEAEVNNNLNATLEQLKMAKTINLESAINQANTDKTTFDNEHPNLVEAYKALKT 170
Db 109 EETPVGTGEATTTTNOANTPATTOSSNTNABELVNO-TSNETTFND-TNTVSSVNS--- 162
QY 171 TLEORANTLEGLST-----AYNOIRNNLDLYNK-----ASSLIT 206
Db 163 --PONSNTAENVSTTQDTSEATPSSNESAPQSDASKNDVVNOAVNTSAPPMRAFSLAA 220
QY 207 KTLPLNGGTLILDSNETTTANKNTNNTLSFINEQ---KTN-ADALSNSFTK----KVIQ 257
Db 221 VAADAPAAGTDI-TNQLTNTVTGIDSGTTVPYHQAGYVKLYNGFSVPNSAVKGTFFKITV 279
QY 258 NNEQSFVGTFTTANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRILANTNSITD 317
Db 280 PKELNLNGVSTAKVPP-----IMAGDQ-----VLANGVIDSD 312
QY 318 VSWIYISLAG-TNTKYQFSFSGNYGPSTCYLYFPYKLVKAADANNVGLQYKLNNGVQQVEF 376
Db 313 GNVIYTFDYVNTKDDYKATLTMPA--YI-----DPENY-----RKTGNV----- 350
QY 377 ATTSANNTANPTPAYDEIKVAKI-----VLSGLRFEGQNTIELSVPTCEGNMKNVAP 429
Db 351 TLAIGIGSTANKTVLDYKEYKYGKFNLSIKGTIDQIDKTNTRYTYVNPSPGDNVIAP 410
QY 430 MI-GNIVLSSNEN 441
Db 411 VLTGNLKXPNTDSN 423

RESULT 9
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.

```

```

; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421.868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match          7.5%; Score 173.5; DB 3; Length 933;
Best Local Similarity 22.5%; Pred. No. 0.0005;
Matches 111; Conservative 66; Mismatches 185; Indels 131; Gaps 23;

QY  2 HYFRNCIFFLIVL-----YGINSS-----PSTQNTVTSREVSVQVLSSESTFYLCPPP 52
DB  9 HAIKKSIGVASVLVGLIFGLLSKKADEASNSVTQSDSASNESKSDSSSVSAAPKT 68
QY  53 VGSTVIRLEFGCMSITKKDANPNNGQTOL--EAARMELTDLINAKAMTLASLDQYAKIEA 110
DB  69 DDTNV-----SDTKTSNTNGETSVAQNPAQQTQSSSTNATT----- 108
QY  111 SLSSAYSEAEVTNNLNLTLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKT 170
DB  109 EETPVGTGEATTTTNOANTPATTTQSSNTNAEELVNO--TSNETTFND--TNTVSSVNS-- 162
QY  171 TLEQRATNLGLST-----AYNQIRNNLVLYNK-----ASSLIT 206
DB  163 --PQNSTNAENVSTQDTSTEAFTPSNESAQSTDAKNQDVNOAVNTSAPRMRAFSLAA 220
QY  207 KTLPLNGGLTLDSEITANKNINNTLSTINEQ-----KTN--ADALSNSEFIK----KVIQ 257
DB  221 VAADAPAAAGTDI--TNOLTNVTGIDSGTTPVPHQAGVYKLVNGYFSPNSAVKGDTEKITV 279
QY  258 NNSQSFVGTETNANQPSNYSFVAFSADVTVPVNYKARVTYVWNGDEPSSRILANTNSITD 317
DB  280 PKELNMGVTSARVPP-----IMAGDQ-----VLANGVIDSD 312
QY  318 VSMIYSLAG--TNTKYOFSEFSNYGPGSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVVEF 376
DB  313 GNVIYFTDVTNKDDVKATLWPA--YI-----DPENV-----KKTGNV----- 350
QY  377 ATSTSANNTANTPTPAVDEIKVAKI-----VLSGLRFQONTIELSVPTGEGNMKNKVP 429
DB  351 TLATGIGSTTANKTVLVDYKYGKYNLSIKGTIDQIDKTNNTYRTIYVNPSPGDNVIAPI 410
QY  430 MI-GNIVLSSNEN 441
DB  411 VLTGNLKPWDSN 423

```

RESULT 10

```

US-08-220-151-4
; Sequence 4, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-220-151-4

Query Match          7.3%; Score 170.5; DB 1; Length 1041;
Best Local Similarity 19.6%; Pred. No. 0.00094;
Matches 117; Conservative 71; Mismatches 188; Indels 221; Gaps 24;

QY  3 YFRRCNCF-----FLIVLY----- 17
DB  22 YFQRCCFFSLLGIAATGSRHGSSGLTFLARYVSVFIWLVFLVGRPVGGSGSTSEQ 81
QY  18 -----GTNSSPSTQNTVTSREVSVQVLSSE-----STFYLCPPVPVSTVIRLE- 61
DB  82 PRRTVATPEVGGTPPKPTTDPDMSDMREALRASQIEANGPSTFYMCPPPSGTVVRLEP 141
QY  62 -----FCMSITKKDANP-----NNGQVQ 80
DB  142 PRACPDYKLGKNFTEGIAVIFKENIAPYFKANIYKNIIMTVWSSGSAVNTNRYDTR 201
QY  81 LEARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAEATNNNLTLEQLKMAKTNL 140
DB  202 VPVKVOEITDLIDRRGMCL-----SKADYVRNNYQFTAFDRDEDPREL 244
QY  141 ESAINQANTDKTTFDNEHPNLVEAYKALKTLEQRATNLEGLSSTAYNOIRNNLVLYNK 200
DB  245 PLKPSKFTPPQSR--GWH-----TYKFKATVYKDVIVSTAWAGSSYQTITRYADRVPI 297
QY  201 ASSLITKTLDPNGGLTLDSEITANKNINNTLSTINEQKTNAD-----ALSNSFIKKV 255
DB  298 PVSEITDITDKFG-----KCSSKATVVRN-NHKVEAFNEKDKPQDMPLTASIKNSVGSKA 351
QY  256 IQNNQSFV-----GTFTNANVQ-----SNYSFVAFSADVTVPVNY----- 291
DB  352 WHTTNETYTKIGAAGFHHSGTSVNCIVEVDARSVPYDSFAISTGDTVHMSPPFFGLRDG 411
QY  292 KYARVTWNGDE-----PSSRILANTNSITDVS-----IYS 323
DB  412 AHVEHTSYSSDRFQIEGYPTDLDTROLGAPVSRNFLTSPHVT--VANNWTPKGRVCT 470
QY  324 LAG-----TNTKYOFSEFSNYGPGSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVVEF 376
DB  471 LAKWEIDEMLRDEYQGSYR-----FTVKIATFISNTS-QFEIN--RIRLGDG 517
QY  377 ATSTSANNTANTPTPAVDEIKVAKIVLSGLRFQONTIELSVPTGEGNMKNKVPMTGN 433
DB  518 ATKEAAE-----AIDRIYKSKYSKTHIQG--TLETYLARG--GFLIAFRPMISN 563

RESULT 11

```

```

US-08-413-118-4
; Sequence 4, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-4

Query Match 7.3%; Score 170.5; DB 1; Length 1041;
Best Local Similarity 19.6%; Pred. No. 0.00094;
Matches 117; Conservative 71; Mismatches 188; Indels 221; Gaps 24;

Qy 3 YFRNCIF-----FLVILY-----17
Db 22 YFRCFFPSSLGLGTAATGSRHNGSSGLRLARVYVFIWLVFLVGPVPGSGSTSEQ 81
Qy 18 -----GTNSSPSTQNTSREVSVSSQLSEER-----STFLCPPPVPGSVIRLE- 61
Db 82 PRRTVATPEVGGTPPKPTTDDTMSDMREALRASQIEANGPSTFIMCPPGSGTVVRLP 141
Qy 62 -----FGCMSITKKDANP-----NNGQTQ 80
Db 142 PRACPDYKLGKNFTGIAVIFKENIAPYKFKANIYYKNIIMTWVWSSGYAVTTNRYTDR 201
Qy 81 LEARMELTDLINAKMTLASLDQYAKIEASLSAYSEATVNNLNATLEQLKMAKTNL 140
Db 202 VPVXVQEIIDLIRRGCL-----SKADYVRNRYQTFADRDDEPREL 244
Qy 141 ESAINQANTDKTTEDNHPNLVEAYKALKTTLEORATNLEGLSSTAYNOIRNNLVLYNK 200
Db 245 PLKPSKNTQSR--GWH-----TYKFKATVYKDVIVSTAWAGSSTQITNRYADRPVI 297
Qy 201 ASSLITKTLDPNGGTLSDNEITANKNNITLSTINEOKTNAD-----ALNSFIKKV 255
Db 298 PVSEITDIDKFG-----KCSSKATVYRN-NHKVEAFNEDKNPQDMLASKYNSVGSKA 351

```

```

Qy 256 IQNNEQSFV-----CTFTNANVQP-----SNYSFVAFSADVTVPVNY----- 291
Db 352 WHTTNETYTKIGAAGFHHSVTSVNCIVEEDARVSPYDVSFAISTGDIHMSPPFGLRDG 411
Qy 292 KYARRTVWNGDE-----PSSRLANTNSITDYSW-----IYS 323
Db 412 ARVHTSYSSDRFQOIEGYPIDLDRQLGAPVSRNRFLETPHYT-VANWNTPKGRVCT 470
Qy 324 LAG-----TNTKYQSFNSYGPSTGYLYFPYKLVRAADANNVGLQYKLNNGNQVVEF 376
Db 471 LAKWREIDEMLRDEYGSYR-----FTVTISATFISNTS-QFEIN--RIRLGDC 517
Qy 377 ATSTSANNTPPAVDKIVKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGN 433
Db 518 ATKEAAE-----AIDRIYKYSKTHIQTG--TLETYLRAG-GFLIAPRPMISN 563

RESULT 12
US-08-473-446-4
; Sequence 4, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-473-446-4

Query Match 7.3%; Score 170.5; DB 3; Length 1041;
Best Local Similarity 19.6%; Pred. No. 0.00094;
Matches 117; Conservative 71; Mismatches 188; Indels 221; Gaps 24;

Qy 3 YFRNCIF-----FLVILY-----17
Db 22 YFRCFFPSSLGLGTAATGSRHNGSSGLRLARVYVFIWLVFLVGPVPGSGSTSEQ 81
Qy 18 -----GTNSSPSTQNTSREVSVSSQLSEER-----STFLCPPPVPGSVIRLE- 61

```

Db 82 PRTVATPEVGGTPPKFTPTDPTDMSDRALRASQIEANGPSTFYMCPPSPGTVVRLRP 141
QY 62 -----FCMSITTKDAMP-----NNGQTQ 80
Db 142 PRACPDYKLGKNTFEGIAVIFKENIAPYKFKANIYVKNIIIMTVWSSGYAVYTNRYTDR 201
QY 81 LEARMELTDLINAKAWTLASQDYAKIEASISAYSEAEVNNNUNATLEQKMAKTWL 140
Db 202 VPVKVQEIITDLIRRGML-----SKADYVRNNYQTFADRDEDPREL 244
QY 141 ESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNOIRNNLVLYNK 200
Db 245 PLKPSFNFQSR-CWH-----TYFKATVYKDVIVSTAGSSYTOITRYADRVDI 297
QY 201 ASSLIKTLDPLNGGLLDSNEITTANKNNINLTSTINEQKTNAD-----ALSNSFIKKV 255
Db 298 PVSEITDITDKFG-----KCSKATVYRN-NHKVEAFNEDKNPQDMPLIASKYNSVGSKA 351
QY 256 IQNNEQSFV-----GTFTNANVP-----SNYSFVAFSADVTPVNY-----291
Db 352 WHITNETYTKIGAAGPHHSGTSVNCIVEVDARSVTPYDSFAISTGDVIMHSPFFGLRDG 411
QY 292 KYARTVWNGDE-----PSSRLANTNSITDVS-----IYS 323
Db 412 ARVHEITSYSDRFQIEGYPIDLDTLRLQLGAPVSRNLETHTVHT-VANNWTFKCGRVCT 470
QY 324 LAG-----TWTKYQFSNPGSTGYLYPKVKAADANNVGLQYKLNNGNQVQVEF 376
Db 471 LAKREIDEMLRDEYQGSYR-----FTVKTISATFISNTS-QFEIN-RIRLGDG 517
QY 377 ATSTSANNTANTPAVDPAIDKIVAKIVLGLRFQNTIELSVPTGEGNMKNKVPMTGN 433
Db 518 ATKEAAE-----AIDRIYKSKYKTHIQTG--TLETYLARG-GFLIAFRPMISN 563

RESULT 13

US-08-836-687B-37
; Sequence 37, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 37
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-37

Query Match
Best Local Similarity 6.6%; Score 154.5; DB 4; Length 682;
Matches 75; Conservative 61; Mismatches 130; Indels 91; Gaps 13;
QY 120 ETVNNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNL 179
Db 38 ELINHNSNALLSSTEGSTDSINLGAQSPAVKSTRTE---LDVTGAATLLQTSAYQK 93
QY 180 EGLSTAYNOIRNNLVLYNKASSLITKLDPLNGGTLDDNEITTANKNNINLTSTINE 239
Db 94 EMKVSQETQVSSE-----FSKRDSTNKEAVPVSKDELLEQSEVVVSTSSQKN-KILDN 148
QY 240 QKTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNTSFVAFSADVTPVNYKARRV- 298
Db 149 KKKRANFTYSPLIKEKPSNKGASGVTDNS-----ASPLSYRAKEVVS 193
QY 299 -----WNGDEPSRLANTNSITDVSWIYSLAGTNTKYQFSFNSYGPS 341
Db 194 LRQPLKNQKVEAQPLLISNSKKASVYTNSHDFWYQW-----DMKY---VTNNGES 243

QY 342 TGYLYFPYKLVKAADANNVGLQYKLNNGNVQOVERATSTAN-----NT 385
Db 244 YA-LVQPSKFI-----SVGI---IDSGIWEHPDLSNLSLGNFYKFNLPKGGFDNEEPDE 293
QY 386 TANPTPAVDPAIDKIVAKIVLGLRFQNTIELSVPTGEGNMKNKVPMTG-NIYLSNEN 441
Db 294 TGNPSDIVD-----KMGHTEVAGQITANGNLGVAPGITYNIVRVFGEN 338
RESULT 14
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6

Query Match
Best Local Similarity 6.3%; Score 146.5; DB 3; Length 1073;
Matches 110; Conservative 89; Mismatches 204; Indels 141; Gaps 25;
QY 17 YGTNSPSTQNTV-----SREVVS-----OLSEESTFYLCPPPGSTVIR 59
Db 542 YETNEAKITTVATDLSQYRESKEYIASLYEKLDRTERNKENENFW-----N 590
QY 60 LEFGCMSITK-----KDANPNQO-TQLEAARMELTDLINAKA-MTLASL---QDYAKIE 109
Db 591 LKFNLLTMLRSHFGSFTDETNGYFTLLDNDFNASELLELTHSNQLLISMTKITEHQSLD 650
QY 110 ASLSAYSEAEVNNLNATLEQKMAKTNLESAINQANTDKTTFDNEHPN-----LVEA 164
Db 651 PALQASRSCAVPNSSLDIYSELKDKSNLLDALEHSLQDISMSQKLGNGISSELIEL 710
QY 165 YKALKTT---LEORATNLEGLSSTAYNOIRNNLVLYNKASSLITKLDPLNGGTLDDSN 221
Db 711 QKDKESYRQLVQELRSYLNQHTHEESQKELMTGVRNDIDALVKTCTTSDADAILSD 770
QY 222 EIT-----TANKNINLTSTINEQ-KTNADAL-----SNSFIKK---VI 256
Db 771 YISDQKSFESKQODLIANICKIVSNFLQEQNESLYTKADILHSHLNDTNSIRKANEIM 830
QY 257 QNNEQSFVGTFTNANVQPSNTSFVAFSADVTPVNYKARRVWNG-----DEPSSRLANT 312
Db 831 NNRSEEFU---RNA-----ASQAEIVGANKERIKQKTVEGNSQLDSSKAIHSNS 877
QY 313 NSITDVSWIYSLA-----GTNTKYQ-----FSPSNYGPSTGYLYFPYKLVKA 354
Db 878 RSMYD-----HCLALAESQKQGVNLEQVQDRLDLQKVEHSEDNTKEHQQL---LDLLES 930
QY 355 ADANNVGL--OYKLNNGNVQOVE---FATSTSANNT---ANPTPAVDEIKV 398
Db 931 LVGNNDNLDSIKYPHTELQKITHVLKGTISLANHNTNELLGLGDESLCNLETTIEDTSL 990
QY 399 AKIVLGLRFQNTI-----ELSVPTGEGNMKNKVPMTG-NIYLSNENNADIKP 448
Db 991 VKLETTGDTSPKRELPAWPSWTRDSSLIKETTNNLSDSKKFFVRETVTSSTQNTNEPDV-- 1048
QY 449 YRRP 452
Db 1049 YDKP 1052

```
RESULT 15
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723.820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-723-820-6

Query Match          6.3%; Score 146.5; DB 4; Length 1073;
Best Local Similarity 20.2%; Pred. No. 0.046;
Matches 110; Conservative 89; Mismatches 204; Indels 141; Gaps 25;

QY 17 YGTNSSPSTQNT-----SREVSV-----QLSEESTFVLCPPPVGSTVIR 59
Db 542 YETNEAKITTVATDLSQYVRESKEYIASLYEKLORTENKKNENFW-----N 590
QY 60 LEFGCMSTK----KDNPNNGQ-TQLEAARMELTDLINAKA-MTLASL----QDYAKIE 109
Db 591 LKFNLTLTMRSGFTDETNGYFTLLNDFNASMEELNTHSNOLLISMTKITEHFQSLD 650
QY 110 ASLSAYSEATVNNLNATLEQLKMAKTNLESAINQANTKTTFDNEHPN-----LVEA 164
Db 651 EALQSARSCAVPNSSLDLIYSELKDSKNSLLDALEHSLQDISMSSQKLGNGISSELI 710
QY 165 YKALKTT---LEQRATNLEGLSSYAYNQIRNNLVLDYNKASSLITKTLDPNGGTLDSN 221
Db 711 QDKMESYRLVQELRSYLNLOHTHEESQKELMGVVRDIDALVKTCTTSLNDADIILSD 770
QY 222 EIT-----TANKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VI 256
Db 771 YISDQKSKFESKQODLIANICKIVSNFLQEQNESLYTKADILHSHLNDTNSIRKANEIM 830
QY 257 QNNEQSFVGTFTNANVQPSNYSFYAFSADVTPVNYKYARRTVWG-----DEPSSRI 312
Db 831 NNRSEEFL---RNA-----ASQAEIVGANKERIQKTVENGSQQLDSKSKAIHSNS 877
QY 313 NSITDVSWIYSLA-----GTNTKIY-----FSPSNYGPSTGYLYFPYKLVKA 354
Db 878 RMYD----HCLALAESQKQGVNLEVQTLDRLLQKVEHSEDNTEKEHQQL---LDLES 930
QY 355 ADANNVGL--QYKLNNGWQOVE-----FATSTANNTT-----ANPTPAVDEIKV 398
Db 931 LVGNNDNLIDSIKTPHTTELQKTDHVLKGTTSANHTNELLGLGDESICNLETTIEDTSL 990
QY 399 AKIVLSGLRFGQNTI-----ELSVPTGEGNANKVAPMIGNIYLSSENENADKIPG 448
Db 991 VKLETTGDTPSKRELPAFPWTRDSSLKETTNLNLDSKKFVRETYTSSNQTNPEPDV-- 1048
QY 449 YRRP 452
Db 1049 YDKP 1052
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:10 ; Search time 37.0487 Seconds
(without alignments)
1953.624 Million cell updates/sec

Title: US-09-147-052-2

Perfect score: 2324

Sequence: 1 MHYFRNCIFLIIVLYGTN.....SSNENNADKIPGYRRPGTFL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19JUN03.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2324	100.0	456	18 AAW36050	Hybrid Marek's dis
2	2011	86.5	1086	18 AAW36051	Hybrid Marek's dis
3	1910	82.2	615	15 AAR63230	Mycoplasma gallise
4	1856	79.9	610	15 AAR63229	Mycoplasma gallise
5	1612	69.4	368	14 AAR44493	Mycoplasma gallise
6	1612	69.4	368	15 AAR63227	Mycoplasma gallise
7	1604	69.0	368	16 AAR76955	Mycoplasma gallise
8	1142	49.1	235	10 AAR93646	Amino acid sequenc
9	1142	49.1	235	11 AAR05081	MG-1 antigen. Myc

10	1142	49.1	261	11 AAR05082	TMG-1 antigen. My
11	1142	49.1	261	16 AAR79911	M.gallisepticum 26
12	1117	48.1	261	10 AAR93959	Amino acid (AA) se
13	807	34.7	661	15 AAR63226	Mycoplasma gallise
14	807	34.7	661	16 AAR79910	M.gallisepticum 66
15	715.5	30.8	647	16 AAW11978	Mycobacterium gall
16	703	30.2	648	15 AAR56973	PMGA 1.2 protein o
17	386.5	16.6	183	10 AAR93649	Amino acid sequenc
18	386.5	16.6	183	11 AAR06439	MG-4 antigen. Myc
19	314	13.5	865	14 AAR30169	Marek's Disease V1
20	284.5	12.2	219	10 AAR93648	Amino acid sequenc
21	284.5	12.2	219	11 AAR06438	MG-3 antigen. Myc
22	187.5	8.1	1095	22 AAG83030	S. epidermidis ope
23	185.5	8.0	10182	23 AAF38314	Staphylococcus epi
24	176.5	7.6	6281	22 AAU37403	Staphylococcus aur
25	174.5	7.5	933	24 ABJ18947	Pathogen specific
26	174.5	7.5	10498	24 ABJ19119	Pathogen specific
27	174	7.5	2086	22 AAU34143	Staphylococcus aur
28	174	7.5	5795	22 AAU37017	Staphylococcus aur
29	173.5	7.5	496	23 AAU73490	S. aureus antigeni
30	173.5	7.5	933	21 AAF58435	Staphylococcus aur
31	173.5	7.5	933	22 AAB69508	Staphylococcus aur
32	173.5	7.5	936	18 AAW89801	Staphylococcus aur
33	170	7.3	520	23 AAE29262	Staphylococcus aur
34	165.5	7.2	2659	23 AAU73489	S. aureus antigeni
35	162	7.0	1029	22 AAU34389	Staphylococcus aur
36	162	7.0	1048	22 AAU37490	Staphylococcus aur
37	161.5	6.9	1237	18 AAW55640	H. pylori ORF 04ep
38	161.5	6.9	1237	20 AAY17187	H. pylori outer me
39	160.5	6.9	1072	23 ABB54963	Lactococcus lactis
40	158	6.8	2434	22 AAU34339	Staphylococcus aur
41	156	6.7	807	21 AAU18311	Plasmodium falcipa
42	155.5	6.7	2478	22 AAU34320	Staphylococcus aur
43	155.5	6.7	2478	22 AAU37374	Staphylococcus aur
44	155.5	6.7	2478	24 ABJ19002	Pathogen specific
45	154.5	6.6	682	17 AAR95273	Nisin nisp gene pr

ALIGNMENTS

RESULT 1

AAW36050 ID AAW36050 standard; Protein; 456 AA.

XX AC AAW36050;

XX DT 15-JUL-1998 (first entry)

XX DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

XX KW antigen; vaccine; poultry.

XX OS Chimeric - Marek's disease gammaherpesvirus.

XX OS Chimeric - Mycoplasma gallisepticum.

XX FH Key Location/Qualifiers

XX FT Region 1..64

XX FT Region /note= "derived from Marek's disease virus gB protein"

XX FT Region 65..456

XX FT Region /note= "derived from M. gallisepticum antigenic protein"

XX PN WO9736924-A1.

XX PD 09-OCT-1997.

XX PF 28-MAR-1997; 97WO-JP01084.

XX PR 29-MAR-1996; 96JP-0103548.

XX PA (JAPG) NIPPON ZEON KK.

PI Saito S, Tsuzaki Y, Yanagida N;
 DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96595.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 456 AA;
 Query Match 100.0%; Score 2324; DB 18; Length 456;
 Best Local Similarity 100.0%; Pred. No. 3.5e-143;
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHYFRRCIFFLIVILYGTSSPTQNTSREVSVSSVQLSEESTFVLCPPVGVSTVIRL 60
 Db 1 MHYFRRCIFFLIVILYGTSSPTQNTSREVSVSSVQLSEESTFVLCPPVGVSTVIRL 60
 QY 61 EFGCMSITKDDANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIESLSAYSEAE 120
 Db 61 EFGCMSITKDDANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIESLSAYSEAE 120
 QY 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEAYKALKTTLEQRATNLE 180
 Db 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEAYKALKTTLEQRATNLE 180
 QY 181 GLSSTAYNOIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSINEQ 240
 Db 181 GLSSTAYNOIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSINEQ 240
 QY 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSEFAVSADVTPVNYKYARTVWN 300
 Db 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSEFAVSADVTPVNYKYARTVWN 300
 QY 301 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSESNYGPSTGYLYFPYKLVKAADANNV 360
 Db 301 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSESNYGPSTGYLYFPYKLVKAADANNV 360
 QY 361 GLOYKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 Db 361 GLOYKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 QY 421 EGNMKNKVPAMIGNIYLSSNENNADKIPGYRRPGTFL 456
 Db 421 EGNMKNKVPAMIGNIYLSSNENNADKIPGYRRPGTFL 456
 RESULT 2
 AAW36051
 ID AAW36051 standard; Protein; 1086 AA.
 XX
 AC AAW36051;
 XX
 DT 15-JUL-1998 (first entry)
 XX
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 XX
 KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 XX antigen; vaccine; poultry.
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX

PH Key Location/Qualifiers
 FT Region 1..672
 FT /note= "derived from Marek's disease virus gB protein"
 FT Region 693..1086
 FT /note= "derived from M. gallisepticum antigen"
 XX
 PN WO9736924-A1.
 XX
 PD 09-OCT-1997.
 XX
 PF 28-MAR-1997; 97WO-JP01084.
 XX
 PR 29-MAR-1996; 96JP-0103548.
 XX
 PA (JAPG) NIPPON ZEON KK.
 XX
 PI Saito S, Tsuzaki Y, Yanagida N;
 XX
 DR WPI; 1997-503046/46.
 DR N-PSDB; AAT96596.
 XX
 PT Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 CC This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 SQ Sequence 1086 AA;
 Query Match 86.5%; Score 2011; DB 18; Length 1086;
 Best Local Similarity 99.7%; Pred. No. 2.6e-122;
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 61 EFGCMSITKDDANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIESLSAYSEAE 120
 Db 691 EFGCMSITKDDANPNNGQTOLEAARMELTDLINAKAMTSLASLDYAKIESLSAYSEAE 750
 QY 121 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEAYKALKTTLEQRATNLE 180
 Db 751 TVNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPLVEAYKALKTTLEQRATNLE 810
 QY 181 GLSSTAYNOIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSINEQ 240
 Db 811 GLSSTAYNOIRNLDVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNINNTLSINEQ 870
 QY 241 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSEFAVSADVTPVNYKYARTVWN 300
 Db 871 KTNADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSEFAVSADVTPVNYKYARTVWN 930
 QY 301 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSESNYGPSTGYLYFPYKLVKAADANNV 360
 Db 931 GDEPSSRIILANTNSITDVSWIYSLAGTNTKYQSFSESNYGPSTGYLYFPYKLVKAADANNV 990
 QY 361 GLOYKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420
 Db 991 GLOYKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 1050
 QY 421 EGNMKNKVPAMIGNIYLSSNENNADKIPGYRRPGTFL 456
 Db 1051 EGNMKNKVPAMIGNIYLSSNENNADKIPGYRRPGTFL 1086
 RESULT 3
 AAR63230
 ID AAR63230 standard; Protein; 615 AA.
 XX


```

AC AAR63230;
XX
XX 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
XX
XX Mycoplasma gallisepticum antigen (UM-67).
XX
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
XX Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
FH Protein 1..615
FT /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A1.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
PR 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
PI Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
DR N-PSDB; AAQ77857.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PT of a live vaccine
XX
XX Claim 4; Page 87-91; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
CC pUM-67 containing an open reading frame was sequenced (AAQ77857). The
CC ORF encodes an antigenic polypeptide (AAR63230). A recombinant avipox
CC virus comprising the coding sequence can be used as a live vaccine to
CC protect against infection by Mycoplasma gallisepticum.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 615 AA;
XX
XX Query Match 82.2%; Score 1910; DB 15; Length 615;
XX Best Local Similarity 98.2%; Pred. No. 4.5e-116;
XX Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
XX
XX 64 CMSTTKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 123
XX
XX 27 CMSTTKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
XX
XX 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
XX
XX 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
XX
XX 184 STAYNQIRNNLVLDYLNKASSLIITKTLDPLNGGTLSDSNEITTANKNNNTLSTNEOKTN 243
XX
XX 147 STAYNQIRNNLVLDYLNKASSLIITKTLDPLNGGTLSDSNEITTANKNNNTLSTNEOKTN 206
XX
XX 244 ADALNSFIKKVIONNESQSFVGTETNANQVPSNFVAFSADVTPVNYKYARRVWNGDE 303
XX
XX 207 ADALNSFIKKVIONNESQSFVGTETNANQVPSNFVAFSADVTPVNYKYARRVWNGDE 266
XX
XX 304 PSSRILANTNSITDVSWYLSLANTKTYQFSFGSTGYLYFPYKLVKAADANNVGLQ 363
XX
XX 267 PSSRILANTNSITDVSWYLSLANTKTYQFSFGSTGYLYFPYKLVKAADANNVGLQ 326
XX
XX 364 YKLANGNVQVFEATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 423

```

```

Db 327 YKLANGNVQVFEATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386
Qy 424 MNKVAPMIGNIYLSNENNADKI 446
Db 387 MNKVAPMIGNIYLSNENNADKI 409

RESULT 4
AAR63229
XX AAR63229 standard; Protein; 610 AA.
XX
XX AAR63229;
XX
XX 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
XX
XX Mycoplasma gallisepticum antigen (UM-66).
XX
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX
XX Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
FH Protein 1..610
FT /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A1.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
PR 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
PI Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
DR N-PSDB; AAQ77856.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PT of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
CC pUM-66 containing an open reading frame was sequenced (AAQ77856). The
CC ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
CC virus comprising the coding sequence can be used as a live vaccine to
CC protect against infection by Mycoplasma gallisepticum.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 610 AA;
XX
XX Query Match 79.9%; Score 1856; DB 15; Length 610;
XX Best Local Similarity 95.5%; Pred. No. 1.5e-112;
XX Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
XX
XX 64 CMSTTKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 123
XX
XX 27 CMSTTKKDANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
XX
XX 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
XX
XX 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 146
XX
XX 184 STAYNQIRNNLVLDYLNKASSLIITKTLDPLNGGTLSDSNEITTANKNNNTLSTNEOKTN 243

```

[illegible]

RESULT 5
AAR44493
ID AAR44493 standard: Protein: 368 AA.

25-MAR-2003	(updated)
16-JUN-1994	(first entry)
XX	
XX	Mycoplasma gallisepticum 40kD antigen.
XX	Vaccine; mycoplasma infection; poultry; fowl.
XX	
XX	Mycoplasma gallisepticum.
OS	

Key	Location/Qualifiers
PH	
FT	
FT	Misc-difference 262
FT	/note= "corresponds to NNN codon in AAQ53419"
FT	
FT	Misc-difference 283
FT	/note= "corresponds to NNN codon in AAQ53419"
XX	
PN	W09324646-A1.

XX	09-DEC-1993.
XX	
XX	28-MAY-1993;
Pf	93WO-JP00715.
XX	
XX	29-MAY-1992;
PR	92JP-Q138819.

PA (JAPG) NIPPON ZEON KK.
PA (SHIO) SHIONOGI & CO LTD.

PI Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;

DR WPI; 1993-405837/50.

XX
PT Mycoplasma gallisepticum antigen and DNA coding for it - useful
PT for vaccination of fowl against mycoplasma infections

PS Claim 2; Page 23-26; 37pp; Japanese.

The sequence coding for the 40kDa antigen was obtained by PCR amplification of *M.gallisepticum* genomic DNA. The antigen reacts with Mycoplasma-immune or Mycoplasma-infected serum and can be used as a vaccine to protect fowl from *M.gallisepticum* infection. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 368 AA;

Query Match	69.4%	Score 1612;	DB 14;	Length 368;
Best Local Similarity	95.0%	Pred. No. 5.6e-97;		
Matches 320; Conservative	6;	Mismatches 11;	Indels 0	

QY	64	CMSITTKDANPNNGOTQLEAAPMELTDLINAKAMTILASQDYAKIEASLSAYSEATVN	123
Db	27	CMSITTKDANPNNGOTQQAARWELTDLINAKARTLASQDYAKIEASLSAYSEATVN	86
QY	124	NNLNATLEQLKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLLEQRATNLEGLS	183
Db	87	NNLNATLEQLKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLLEQRATNLEGLA	146
QY	184	STAYNOIRNNLVLDYKNKASSLTKTLDPLNGTGLDLSNEITTKANKNINTLSTINEQKN	243
Db	147	STAYNOIRNNLVLDYNNKASSLTKTLDPLNGMLDLSNEITTVNRNINTLSTINEQKN	206
QY	244	ADALNSFIKKVIQNNQESFVGTFTNANVPNSYFVAFSADVPVNYKYARRTVYNGDE	303
Db	207	ADALNSFIKKVIQNNQESFVGTFTNANVPNSYFVAFSADVPVNYKYARRTVYNGDE	266
QY	304	PSSRILANTNSTDVSWYISLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ	363
Db	267	PSSRILANTNSTDVSXYISLAGTNTKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ	326
QY	364	YKLNGNVOQVEFATSTANNNTANPTPAVDIEKVA 400	
Db	327	YKLNGNVOQVEFATSTANNNTANPTQOOLMKLKLK 363	

RESULT 6	
AAR63227	
ID	AAR63227 standard; Protein; 368 AA.
XX	
XX	
AC	AAR63227;
XX	
XX	
DT	25-MAR-2003 (updated)
DT	23-JUN-1995 (first entry)
XX	
XX	
DE	Mycoplasma gallisepticum 40kD antigen.
XX	
KW	recombinant avipox virus; live vaccine; mycoplasma 40kD antigen;
KW	TFM-1.

AA			
OS	Mycoplasma gallisepticum.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 262	/note=	"corresponds to a NNN codon"
FT	Misc-difference 283	/note=	"corresponds to a NNN codon"
FT			
FT			

XX	WO9423019-A1.	
PN		
XX		
XX		
PD	13-OCT-1994.	
XX		
XX		
PF	31-MAR-1994; 94WO-JP00541.	
XX		
XX		
PR	31-MAR-1993; 93JP-0074139.	
PR	30-SEP-1993; 93JP-0345625.	
XX		
XX		
PA	(JAPG) NIPPON ZEON KK.	
PA	(SHIO) SHIONOGI & CO LTD.	
XX		
PI	Aoyama S, Funato H, Iritani Y,	
PI	Saeki S, Saitos, Takahashi K;	
XX		
XX		
DR	WPI; 1994-333181/41.	
DR	N-PSDB; AAQ77854.	
DR		

AA Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PT of a live vaccine
PT

PS Claim 4; Page 71-74; 123pp; Japanese. 2x

CC The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding

CC for the 40kD antigen of Mycoplasma gallisepticum under the control
 CC of a synthetic promoter. A 1300 bp restriction fragment containing
 CC the promoter-ORF sequence was excised and was used in the
 CC construction of plasmid pNZ7929-R2. This in turn was involved in the
 CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle Disease Virus haemagglutinin, neuraminidase and FpV
 CC sequences. The recombinant avipox virus is useful as a live vaccine
 CC to protect against infection by Mycoplasma gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX

SQ Sequence 368 AA;

Query Match 69.4%; Score 1612; DB 15; Length 368;
 Best Local Similarity 95.0%; Pred. No. 5.6e-97;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 64 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVYN 123
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKARTLASLDYAKIEASLSAYSEAEVYN 86
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 QY 184 STAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTNANKNINNTLSTINEQKTN 243
 DB 147 STAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTVNRNINNTLSTINEQKTN 206
 QY 244 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 303
 DB 207 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
 QY 304 PSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 363
 DB 267 PSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 326
 QY 364 YKLNGNVQOQVEFATSTANNTTANPTPAVDKIKVAK 400
 DB 327 YKLNGNVQOQVEFATSTANNTTANPTQOQLMKLKLK 363

RESULT 7

AAR76955
 ID AAR76955 standard; Protein; 368 AA.

XX AAR76955;

XX 25-MAR-2003 (updated)
 DT 12-MAR-1996 (first entry)

XX Mycoplasma gallisepticum antigenic protein TTM-1.

DE Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Misc-difference 262 /note= "any amino acid"

FT Misc-difference 283 /note= "any amino acid"

XX JP07133295-A.

XX 23-MAY-1995.

XX 27-AUG-1993; 93JP-0213102.

XX 27-AUG-1993; 93JP-0213102.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI & CO LTD.

PA

XX WPI; 1995-220782/29.
 DR N-PSDB; AAQ94711.

XX A new antigenic protein which reacts with Mycoplasma gallisepticum -
 PT is useful in a component vaccine for use against poultry infected
 PT with M. gallisepticum.

XX Claim 6; Figs 5-6; 33pp; Japanese.

XX AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
 CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
 CC infectious diseases in poultry, and as a diagnostic agent for
 CC M. gallisepticum.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX

SQ Sequence 368 AA;

Query Match 69.0%; Score 1604; DB 16; Length 368;
 Best Local Similarity 94.4%; Pred. No. 1.8e-96;
 Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 64 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVYN 123
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKARTLASLDYAKIEASLSAYSEAEVYN 86

QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTDFNEHPNLVEAYKALKTTLEQRATNLEGLA 146

QY 184 STAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTNANKNINNTLSTINEQKTN 243
 DB 147 STAYNQIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTVNRNINNTLSTINEQKTN 206

QY 244 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 303
 DB 207 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 266

QY 304 PSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 363
 DB 267 PSSRILANTNSITDVSIIYSLAGTNTKYQFSFSGTGYLYFPYKLVKAADANNVGLQ 326

QY 364 YKLNGNVQOQVEFATSTANNTTANPTPAVDKIKVAK 400

DB 327 YKLNGNVQOQVEFATSTANNTTANPTQOQLMKLKLK 363

RESULT 8

AAP93646
 ID AAP93646 standard; protein; 235 AA.

XX AAP93646;

XX 25-MAR-2003 (updated)

DT 11-MAY-1990 (first entry)

XX Amino acid sequence of Mycoplasma gallisepticum (MGI) polypeptide.

DE Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0305441.

XX 31-MAY-1989; 89EP-0305441.

XX 02-JUN-1988; 88JP-0136343.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI SEIYAKU KK.

PA (JAPG) NIPPON ZEON KK.
 XX (SHIO) SHIONOGI SEIYAKU KK.
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
 XX WPI; 1989-358393/49.
 DR N-PSDB; AAN92568.
 XX
 PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
 PT vaccines.
 XX
 PS Disclosure; Fig.1a; 3lpp; English.
 XX
 CC This amino acid sequence of MG1 is encoded by M1 DNA and elicits an
 CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
 CC (Updated on 25-MAR-2003 to correct Pf field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct FI field.)
 XX
 SQ Sequence 235 AA;

Query Match 49.1%; Score 1142; DB 10; Length 235;
 Best Local Similarity 97.0%; Pred. No. 1.2e-66;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 64 CMSITKDDPNNGQTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 123
 DB 1 CMSITKDDPNNGQTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 120
 QY 184 STAYNQIRNRLVDLYNNKASSLIKTLDPLNGGTLSDSNEITTANKNNINNTLSINEOKTN 243
 DB 121 STAYNQIRNRLVDLYNNKASSLIKTLDPLNGGTLSDSNEITTANKNNINNTLSINEOKTN 180
 QY 244 ADALNSFIKKVQNNNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 298
 DB 181 ADALNSFIKKVQNNNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 235

RESULT 9
 AAR05081
 ID AAR05081 standard; protein; 235 AA.
 AC AAR05081;
 XX
 XX 10-MAR-2003 (updated)
 DT 08-OCT-1990 (first entry)
 XX
 XX MG-1 antigen.
 DE
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX
 OS Mycoplasma gallisepticum.
 XX
 PN JP02111795-A.
 XX
 PD 24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX
 DR WPI; 1990-169109/22.
 DR N-PSDB; AAQ04686.

XX
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX
 PS Claim 2; Fig 1a; 20pp; Japanese.
 XX
 CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081-2 and AAR06437-41.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 XX
 SQ Sequence 235 AA;

Query Match 49.1%; Score 1142; DB 11; Length 235;
 Best Local Similarity 97.0%; Pred. No. 1.2e-66;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 64 CMSITKDDPNNGQTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 123
 DB 1 CMSITKDDPNNGQTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 120
 QY 184 STAYNQIRNRLVDLYNNKASSLIKTLDPLNGGTLSDSNEITTANKNNINNTLSINEOKTN 243
 DB 121 STAYNQIRNRLVDLYNNKASSLIKTLDPLNGGTLSDSNEITTANKNNINNTLSINEOKTN 180
 QY 244 ADALNSFIKKVQNNNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 298
 DB 181 ADALNSFIKKVQNNNEQSFVGTFTNANQPSNYSFVAFSADVTPVNYKYARRTV 235

RESULT 10
 AAR05082
 ID AAR05082 standard; protein; 261 AA.
 AC AAR05082;
 XX
 XX 10-MAR-2003 (updated)
 DT 08-OCT-1990 (first entry)
 XX
 XX TMG-1 antigen.
 DE
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX
 OS Mycoplasma gallisepticum.
 XX
 PN JP02111795-A.
 XX
 PD 24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX
 DR WPI; 1990-169109/22.
 DR N-PSDB; AAQ04687.
 XX
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX
 PS Claim 2; Fig 2; 20pp; Japanese.

CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 CC
 XX Sequence 261 AA;

Query Match 49.1%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 1.3e-66;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 64 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 123
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKARTILASLDYAKIEASLSAYSEAEVTN 86
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 QY 184 STAYNOIRNNLVLYNKASSLIKTLDPLNGGTLDSNEITTANKNNNTLSTINEQKTN 243
 DB 147 STAYNOIRNNLVLYNNASSLIKTLDPLNGGMLDSNEITTVNRNNTLSTINEQKTN 206
 QY 244 ADALSNFSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 298
 DB 207 ADALSNFSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 261

RESULT 11

AAR79911
 ID AAR79911 standard; Protein; 261 AA.

XX AC AAR79911;

XX 19-JUL-1996 (first entry)

XX M.gallisepticum 261 amino acid protein.

XX Detection; probe; primer; PCR; amplification; secretion; lung;
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.
 XX Mycoplasma gallisepticum.

XX JP07236498-A.

XX 12-SEP-1995.

XX 25-FEB-1994; 94JP-0052764.

XX 25-FEB-1994; 94JP-0052764.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI & CO LTD.

XX WPI; 1995-347462/45.

XX N-PSDB; AAT04076.

XX Detection of Mycoplasma gallisepticum - for the quick detection,
 PT i.e. within one day, of avian chronic respiratory diseases

XX Claim 3; Page 10-11; 11pp; Japanese.

XX This is the amino acid sequence of a 261 amino acid protein encoded
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid
 CC protein) can be used to detect M.gallisepticum using probes based on
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
 CC complement of bases 893-919, 1908-1934 and the complement of bases
 CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
 CC encoding this protein. The method using these sequences is faster i.e. is

CC able to detect M.gallisepticum, which causes avian chronic respiratory
 CC diseases, within one day, from avian secretions, washings from the lung,
 CC respiratory tract, nasal cavity, etc.

SQ Sequence 261 AA;

Query Match 49.1%; Score 1142; DB 16; Length 261;
 Best Local Similarity 97.0%; Pred. No. 1.3e-66;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 64 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 123
 DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKARTILASLDYAKIEASLSAYSEAEVTN 86
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183
 DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 QY 184 STAYNOIRNNLVLYNKASSLIKTLDPLNGGTLDSNEITTANKNNNTLSTINEQKTN 243
 DB 147 STAYNOIRNNLVLYNNASSLIKTLDPLNGGMLDSNEITTVNRNNTLSTINEQKTN 206
 QY 244 ADALSNFSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 298
 DB 207 ADALSNFSFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV 261

RESULT 12

AAP93959

ID AAP93959 standard; protein; 261 AA.

XX AC AAP93959;

XX 25-MAR-2003 (updated)

DT 11-MAY-1990 (first entry)

XX Amino acid (AA) sequence of TMG-1 polypeptide.

XX Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0305441.

XX 31-MAY-1989; 89EP-0305441.

XX 02-JUN-1988; 88JP-0136343.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI SEIYAKU KK.

XX (JAPG) NIPPON ZEON KK.

XX (SHIO) SHIONOGI SEIYAKU KK.

XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX WPI; 1989-358393/49.

XX N-PSDB; AAN92574.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
 PT vaccines.

XX Disclosure; Fig.2; 31pp; English.

XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
 CC same sequence as that of a polypeptide expressed in Mycoplasma
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted
 CC into a recombinant vector used to transform a host the antigen protein
 CC produced can be used as a vaccine to prevent and diagnose MG infection.
 CC (updated on 25-MAR-2003 to correct PF field.)

```

CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 261 AA;

Query Match 48.1%; Score 1117; DB 10; Length 261;
Best Local Similarity 96.2%; Pred. No. 5.7e-65;
Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;

QY 64 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASIQDYAKIEASLSAYSEATVN 123
DB 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKARTILASIQDYAKIEASLSAYSEATVN 86
QY 124 NNLNATLEOLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTILEQRTATNLEGLS 183
DB 87 NNLNATLEOLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTILEQRTATNLEGLA 146
QY 184 STAYNQIRNNLDYNKASSLTKTLDPLNGGTLDSNEITVANKNNINNTLTINE-OKT 242
DB 147 STAYNQIRNNLDYNKASSLTKTLDPLNGGMLDSNEITVNRN-NNTLTINEQOQKT 205
QY 243 NADALSNSFIKKVIONNEQSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRTV 298
DB 206 NADALSNSFIKKVIONNEQSFVGTETNANVQPSNYSFVAFSADVTPVNYKYARRTV 261

RESULT 13
AAR63226
ID AAR63226 standard; Protein; 661 AA.
XX AC AAR63226;
XX DT 25-MAR-2003 (updated)
XX DT 23-JUN-1995 (first entry)
XX DE Mycoplasma gallisepticum antigen (UM-81).
XX KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX OS Mycoplasma gallisepticum.
XX FH Key Location/Qualifiers
XX FT Protein 1..661
XX FT /note= "Trp residues correspond to TGA codons"
XX PN W09423019-A1.
XX PD 13-OCT-1994.
XX PF 31-MAR-1994; 94WO-JP00541.
XX PR 31-MAR-1993; 93JP-0074139.
XX PR 30-SEP-1993; 93JP-0245625.
XX XX (JAPG ) NIPPON ZEON KK.
XX PA (SHIO ) SHIONOGI & CO LTD.
XX PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX PI Saeki S, Saitos, Takahashi K;
XX XX WPI; 1994-333181/41.
XX DR N-PSDB; AAQ77853.
XX XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX PT exhibiting antigenicity of mycoplasma, useful for the production
XX PT of a live vaccine
XX PS Claim 4; Page 61-65; 123pp; Japanese.
XX PS
XX CC A restriction fragment of the insert of M.gallisepticum genomic clone
XX CC PUM-81 containing an open reading frame was sequenced (AAQ77853). The
XX CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox
XX CC virus comprising the coding sequence can be used as a live vaccine to

```

```

CC protect against infection by Mycoplasma gallisepticum.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 661 AA;

Query Match 34.7%; Score 807; DB 15; Length 661;
Best Local Similarity 41.6%; Pred. No. 2.9e-44;
Matches 188; Conservative 64; Mismatches 128; Indels 72; Gaps 13;

QY 51 PPGVSTVIRLEGGCSIRKKDANPNNGQTQLEAARMELTD-----LINAKMTILASIQ 103
DB 35 PPSGG-----NMNGGNTNPDSGOGMMAAKELADAKAALTTTLINGETANLASYE 84
QY 104 DYAKIEASLSAYSEATVNNLNATLEOLKAKTNLESAINQANTDKTTFDNEHPNLVE 163
DB 85 DYAKIKSELTSAYETAKAVSAKTGATLNEVNEAKTLDNAIKKASAKNDPDAQHGSUVE 144
QY 164 AYKALKTILEQRTATNLEGLSAYNOIRNNLDYNKASSLTKTLDPLNGGTLDSNEI 223
DB 145 AYNNLKETLKEEKTLDLSLANENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPVMSV 203
QY 224 TTANKNINNTLTINEQKTNADALSNSFIKKVIONNEQSFVGTETNANVQPSNYSFVAFS 283
DB 204 TQANDITNATSLRLTAWKQNDNLANSFIKQSLVKNLNRVDVANNQE-OPANYSFVGFS 262
QY 284 ADVTVPVNYKYARRTVNMGDE-----PSRILANTNSITDVSWIYSLAGTNTKYQFSFN 337
DB 263 VNVDTPNWNAQKQVWASENTPLATTPAEDATQQAASLTDSWIYSLNGAEAKYTLSEFY 322
QY 338 YG-PSTGYLYPPYKLVKAADANNVGLYKLNNGVQQVEF-----ATSTSAN--NTTA 387
DB 323 FGAERTAYLYPPYKLVKTSQ--NVGLYKLNNGDTKQINFVQTPASGSSDVAANEETMA 380
QY 388 NP-----TPAYDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPIMIGNIYLS-- 438
DB 381 SPAEQSAPTVDIKIAKVALSNLAFNSNTIEFSVPTG-----KAAPMIGNMYLTSSNSE 435
QY 439 -----NENN-----ADKIPGY 449.
DB 436 VNKNKIYDDLFGNSFNENNPTAVTVDLLGY 467

RESULT 14
AAR79910
ID AAR79910 standard; Protein; 661 AA.
XX AC AAR79910;
XX DT 19-JUL-1996 (first entry)
XX DE M.gallisepticum 661 amino acid protein.
XX KW Detection; probe; primer; PCR; amplification; secretion; lung;
XX KW avian chronic respiratory disease; respiratory tract; nasal cavity.
XX OS Mycoplasma gallisepticum.
XX FH Key Location/Qualifiers
XX FT Misc-difference 220 /note= "encoded by TGA"
XX FT Misc-difference 270 /note= "encoded by TGA"
XX FT Misc-difference 305 /note= "encoded by TGA"
XX FT Misc-difference 581 /note= "encoded by TGA"
XX XX JP07236498-A.
XX XX 12-SEP-1995.
XX XX 25-FEB-1994; 94JJP-0052764.
XX

```

```

PR 25-FEB-1994; 94JP-0052764.
XX (JAPC ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
XX WPI; 1995-347462/45.
DR N-PSDB; AAT04075.
XX
XX Detection of Mycoplasma gallisepticum - for the quick detection,
PT i.e. within one day, of avian chronic respiratory diseases
XX
XX Claim 2; Page 7-10; 11pp; Japanese.
XX
XX This is the amino acid sequence of a 661 amino acid protein encoded
CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
CC sequence and the sequence of AAT04076 (encoding a 261 amino acid
CC protein) can be used to detect M.gallisepticum using probes based on
CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
CC complement of bases 893-919, 1908-1934 and the complement of bases
CC 2184-2210 of the sequence, and a probe based on nucleotides 718-41 of
CC AAT04076. The method using these sequences is faster i.e. is able to
CC detect M.gallisepticum, which causes avian chronic respiratory diseases,
CC within one day, from avian secretions, washings from the lung,
CC respiratory tract, nasal cavity, etc.
XX
XX Sequence 661 AA;
XX
Query Match 34.7%; Score 807; DB 16; Length 661;
Best Local Similarity 41.6%; Pred. No. 2.9e-44;
Matches 188; Conservative 64; Mismatches 128; Indels 72; Gaps 13;
QY 51 PPVGSTVIRLEPGCMSTYKDDPNNGQTOLEAARMELTD-----LINAKAMTLASIQ 103
DB 35 PPSGG-----NMGGNTNPSDGGQGMNAAAKELADAKAALTTLINGETANLASYE 84
QY 104 DYAKTEASISAYSAETVNNLNATLEQKMAKTNLESAINQANTDKTTDNEHPNLVE 163
DB 85 DYAKIKSELTSAIETAKAVSAKTGATLNEVEAKTTLDAIKKAASAKNDFDAQHGSLSVE 144
QY 164 AYKALKTTLEQRATNLEGLSTAYNOIRNNLDVLYNKASSLTKTKLDPLNGTGLDLSNEI 223
DB 145 AYNNLKETLKEKTKNLDLSANENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPVMSV 203
QY 224 TTANKNINNTLSTINEQTNADALSNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFS 283
DB 204 TOANQDITNATSLTAWKONANLANSFKIKOSLVKNLNRDVRDANNQE-OPANYSFVGF 262
QY 284 ADVTPVNTKYARTVWNGDE-----PSSRILANTNSITDVSWIYSLAGTNTKYQFSFN 337
DB 263 VNVDPNPNFAQRKYWASENTPLATTPAEDATQQAASLTDVSWIYSLNGAEAKYTLSEY 322
QY 338 YG-PSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVEF-----ATTSAN--NTTA 387
DB 323 FGAECTAYLYFPYKLVKTSQ--NVGLQYKLNNGDTFKQINFVQTPASGSDVAANEETMA 380
QY 388 NP-----TPADEVKIAKIVLSGLRFGONTLELSPVTGEGNNKVPAMIGNIYLS-- 438
DB 381 SPAEQSAPTDDIKIAKVALSNLKNFNTSEFVSPTG-----KAAPMIGNMYLTSSNSE 435
QY 439 -----NENN-----ADKIPY 449
DB 436 VNKNIYDDLFNGSNENNENTTAVTVDLKGY 467
RESULT 15
AAW11978
ID AAW11978 standard; Protein; 647 AA.
XX
XX AAW11978;
AC
XX
XX 23-APR-1997 (first entry)
XX
XX Mycobacterium gallisepticum pmGAL.2 adhesin.
DE

```

```

XX Adhesin; pmGA; mycoplasma; diagnosis; vaccine; vector;
KW respiratory disease; poultry; haemagglutinin.
XX
XX Mycoplasma gallisepticum strain S6.
OS
XX Key Location/Qualifiers
FT 1..25
FT /label= sig_peptide
FT /note= "the signal peptide shows homology to
FT the pmGAL.3 signal peptide"
XX
XX CA2135330-A.
XX
XX 11-MAY-1995.
XX
XX 08-NOV-1994; 94CA-2135330.
XX
XX 20-APR-1994; 94US-0230312.
XX 10-NOV-1993; 93AU-0050593.
XX (BROW/) BROWNING G F.
XX
XX Browning GF, Glew MD, Markham PF, Walker ID, Whithear KG;
XX
XX WPI; 1995-241027/32.
XX N-PSDB; AAT51531.
XX
XX New promoter region from a Mycoplasma gallisepticum adhesin gene -
XX useful when coupled to foreign antigen gene, for prodn. of
XX multivalent live vaccines, also new probes for detecting Mycoplasma
XX and manipulating its genome
XX
XX Disclosure; Fig 3; 81pp; English.
XX
XX Adhesin pmGAL.2 (AAW11978) and adhesin pmGAL.3 fragment (AAW11979)
XX are products of gene sequences (see also AAT51531) isolated from
XX Mycoplasma gallisepticum. DNA constructs incorporating the
XX promoter and/or signal sequences of the pmGA genes can be used
XX in the prodn. of multivalent live vaccines. The signal peptide
XX sequence is utilised where attachment of an exogenous antigen
XX gene to the mycoplasma cell membrane is required.
XX
XX Sequence 647 AA;
XX
Query Match 30.8%; Score 715.5; DB 16; Length 647;
Best Local Similarity 39.2%; Pred. No. 2.6e-38;
Matches 174; Conservative 63; Mismatches 134; Indels 73; Gaps 13;
QY 53 VGSTVIRLEFGCMSTTK-----DANPNNGQ-----TOLEAARMELTD 90
DB 15 IGSFYMLAAASCTTPTNPTNPPSGMNGGDTNPGDQGMNAAASQELAAARMGLTT 74
QY 91 LINAKAMTLASIQDYAKTEASISAYSAETVNNLNATLEQKMAKTNLESAINQANTD 150
DB 75 IPDSKAKNLGLVYDKTKQNTLTAKDAKTDLNDSSTQKLNKAKTLETAIRTAAYS 134
QY 151 KTFPNEHPNLVEAYKALKTKTLEQRATNLEGLSTAYNOIRNNLDVLYNKASSLTKTLD 210
DB 135 KQTFDEQHAELVKVYKELTKTTLSTNETATLAPYADAQYAGIKMHLSCLYDAGKAITTKLE 194
QY 211 PLNGTGLDLSNEITANKNINNTL--STINEQTN-----ADALSNSFKKVIQNEQSFV 264
DB 195 PVEGDP-LTASAVMMANTKIVEAIKDEVLPNPKENATKLADSLSSIVKKITGVEE---- 249
QY 265 GTFTNANVQPSNYSFVAFSADVTVP-----NKKYARTVW-NGDEPSSRILANT--- 312
DB 250 ---AHNAQAPANYSFVGYKRWYTELLDKQVFPNMDYAEIRTFITNSDEP--RSISNTPAD 304
QY 313 -----NSITDVSWIYSLAGTNTKYQFSFNYPSTGYLYFPYKLVKAADANNVGLQYKLN 367
DB 305 GQTMQAQPLSNVSWIYSLAGTGAKYKLEFTTYGYPSTGYLYFPYKLVNTSDQVKLGLEYKLN 364

```

QY	368	NGNVQQVEFATSTSA-----NNTTANTPTPAVDEIKAVKIVLSGLRFGQNTIELSVPTGEG	422
Db	365	D-----ATKPSAITFGSDQTMGKITPTVDINVAKTILANLFGSNKFIEFSVPA---	413
QY	423	NNMKVAPMIGNIYLSSNNENADKI	446
Db	414	--EKVSPMIGNMYLSSSPNNWNKI	435

Search completed: October 8, 2003, 17:02:46
Job time : 39.0487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2003, 16:55:10 ; Search time 54.5982 Seconds
(without alignments)
1953.624 Million cell updates/sec

Title: US-09-147-052-4_COPY_1_672
Perfect score: 3510
Sequence: 1 MHYFRNCIFLVLVYGTN.....DVGVLDAEVARRNQHLK 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	3510	100.0	865	14	AAR30169	Marek's Disease V1
2	3503	99.8	1086	18	AAW36051	Hybrid Marek's dis
3	1898.5	54.1	879	16	AAR77024	Canine herpesvirus
4	1726	49.2	868	13	AAW22615	Varicella-zoster v
5	1719.5	49.0	913	19	AAW68404	Aujeszky's disease
6	1715.5	48.9	845	6	AAP50035	N-terminal sequenc
7	1699	48.4	891	17	AAW92746	B virus gB glycopr
8	1699	48.4	891	19	AAW70293	Simian herpesvirus
9	1695.5	48.3	933	13	AAR27807	Bovine herpes viru

10	1695.5	48.3	933	14	AAW41343	Bovine herpesvirus
11	1693	48.2	904	23	AAE17812	Herpes simplex vir
12	1692	48.2	854	19	AAW72113	HSV-2 strain SB5 C
13	1692	48.2	904	14	AAW41778	Glycoprotein B (gB
14	1692	48.2	904	17	AAW00376	HSV-2 glycoprotein
15	1692	48.2	904	18	AAW34552	Herpes simplex vir
16	1692	48.2	904	19	AAW72193	HSV-2 strain SB5 C
17	1692	48.2	904	22	AAW74442	Herpes simplex vir
18	1683.5	48.0	943	21	AAW32470	DNA encoding felin
19	1683.5	48.0	943	22	AAW50113	Feline herpesvirus
20	1683	47.9	904	9	AAP80914	Sequence of Herpes
21	1679	47.8	854	8	AAP70347	Varicella-zoster v
22	1679	47.8	885	17	AAW92747	SA8 virus gB glyco
23	1676	47.7	907	8	AAW71136	Herpes Simplex vir
24	1663	47.4	795	19	AAW72062	HSV-2 strain SB5 C
25	1663	47.4	904	17	AAW00375	HSV-1 glycoprotein
26	1663	47.4	904	22	AAW74441	Herpes simplex vir
27	1661	47.3	903	8	AAP70426	Recombinant herpes
28	1660	47.3	694	12	AAW14666	Truncated HSVgB po
29	1660	47.3	904	12	AAW14665	HSVgB polypeptide.
30	1660	47.3	973	12	AAW14680	HSV surface antige
31	1658	47.2	903	7	AAP60244	Herpes simplex vir
32	1657	47.2	903	8	AAP71135	Herpes Simplex Vir
33	1654	47.1	904	14	AAW41779	Glycoprotein B (gB
34	1654	47.1	905	9	AAP80315	Sequence of Herpes
35	1652	47.1	973	8	AAP70769	Glycoprotein B of
36	1648	47.0	904	18	AAW34553	Herpes simplex vir
37	1546.5	44.1	928	16	AAW77399	BHV1 gI glycoprote
38	1533.5	43.7	903	6	AAP50312	Herpes simplex vir
39	1294	36.9	932	19	AAW44947	Bovine herpesvirus
40	1293	36.9	873	16	AAW79459	Infectious laryngo
41	1163	33.1	825	13	AAW22237	Sequence of the 'g
42	773.5	22.0	400	8	AAP70654	Sequence encoded b
43	761	21.7	907	15	AAW63462	Human cytomegalovi
44	761	21.7	907	18	AAW27274	Human cytomegalovi
45	754	21.5	906	10	AAW91047	gB envelope protei

ALIGNMENTS

RESULT 1	
AAR30169	AAR30169 standard; Protein; 865 AA.
ID	XX
AC	XX
DT	XX
DT	XX
DT	XX
DE	XX
KW	XX
KW	XX
KW	XX
OS	XX
PN	XX
XX	XX
XX	XX
PD	XX
PF	XX
XX	XX
PR	XX
PA	XX
PA	XX
XX	XX
PI	XX
XX	XX
DR	XX
XX	XX

PT Recombinant fowl pox virus contg. Marek's disease virus antigen
PT gene - used to produce cell-free vaccine against Marek's disease
PT virus
XX
PS Example 2; Page 15-19; 30pp; English.
XX
CC The MDV gB of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-SalI
CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
CC from K3 fragment were ligated with EcoRI, SalI digested pUC18. The
CC sequence of the putative MDV gB was determined by sequencing a set
CC of deletion mutants. The nucleotide and amino acid sequences were
CC found to be identical with the published sequences of the gB of
CC RB18 strain of MDV (Ross et al., J. Gen. Virol., 70:1789-1994, 1988).
CC A fragment contg. the entire coding region of MDV gB was inserted
CC into pNZ1729R (see AQ34774-Q34778) to produce a recombinant FPV/MDVgB
CC virus for immunising chickens. See also AQ34780-Q34781.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 865 AA;
Query Match 100.0%; Score 3510; DB 14; Length 865;
Best Local Similarity 100.0%; Pred. No. 1.6e-298;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHYFRRNCIFFLIVILYGTNSPSTQNTVTSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRRNCIFFLIVILYGTNSPSTQNTVTSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
QY 61 EPPKPCPEPKKATGEGGIAILFKENISPYKFKVLYYKNIQTWTGTYRQITNRYT 120
DB 61 EPPKPCPEPKKATGEGGIAILFKENISPYKFKVLYYKNIQTWTGTYRQITNRYT 120
QY 121 DRTPVSTEEITDLDGKGRSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
DB 121 DRTPVSTEEITDLDGKGRSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
QY 181 HTTNETYTVWGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVWGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPFYGLSPPE 240

RESULT 2
AAW36051 AAW36051 standard; Protein: 1086 AA.
XX
AC AAW36051;
XX
DT 15-JUL-1998 (first entry)
XX
DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX
KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW antigen; vaccine; poultry.
XX
OS Chimeric - Marek's disease gammaherpesvirus.
OS Chimeric - Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Region 1..672
FT /note= "derived from Marek's disease virus gB protein"
FT Region 693..1086
FT /note= "derived from M. gallisepticum antigen"
XX
PN W09736924-A1.
XX
PD 09-OCT-1997.
XX
PF 28-MAR-1997; 97WO-JP01084.
XX
PR 29-MAR-1996; 96JP-0103548.
XX
PA (JAPG) NIPPON ZEON KK.
XX
PI Saito S, Tsuzaki Y, Yanagida N;
XX
DR WPI; 1997-503046/46.
DR N-PSDB; AAT96596.
XX
PT Fusion protein comprising herpes virus outer membrane protein and
PT antigenic polypeptide - for prevention of infection by Mycoplasma
PT gallisepticum, especially in poultry
XX
PS Disclosure; Page 22-30; 51pp; Japanese.
XX
CC This sequence represents the chimeric protein 40 K-C which comprises a
CC fragment of the Marek's disease virus outer membrane protein gB fused
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
CC protein can be used in recombinant live vaccines for prevention of
CC infection by Mycoplasma gallisepticum, especially as the outer membrane
CC protein shows antigenicity in poultry.
XX
SQ Sequence 1086 AA;
Query Match 99.8%; Score 3503; DB 18; Length 1086;
Best Local Similarity 99.7%; Pred. No. 9.7e-298;
Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MHYFRRNCIFFLIVILYGTNSPSTQNTVTSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
DB 1 MHYFRRNCIFFLIVILYGTNSPSTQNTVTSREVSVSSVQLSEESTFYLCPPVGVSTVIRL 60
QY 61 EPPKPCPEPKKATGEGGIAILFKENISPYKFKVLYYKNIQTWTGTYRQITNRYT 120
DB 61 EPPKPCPEPKKATGEGGIAILFKENISPYKFKVLYYKNIQTWTGTYRQITNRYT 120
QY 121 DRTPVSTEEITDLDGKGRSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
DB 121 DRTPVSTEEITDLDGKGRSSKARYLRNNVYVEAFDRDAGEKQVLLKPSKFNTPESRAW 180
QY 181 HTTNETYTVWGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPFYGLSPPE 240
DB 181 HTTNETYTVWGSPWYRTGTSVNCIVEEMDARSFPYSYFAMANGDIANISPFYGLSPPE 240

XX	AC	AAR22615;
XX	DT	25-MAR-2003 (updated)
XX	DT	21-OCT-1992 (first entry)
XX	DE	Varicella-zoster virus polypeptide.
XX	KW	VZV; protective; immunogenic; gB glycoproteins; chickenpox; herpesvirus.
XX	OS	Varicella zoster virus.
XX	PN	EP482671-A.
XX	PD	29-APR-1992.
XX	PF	31-JUL-1986; 91EP-0120701.
XX	PR	02-AUG-1985; 85US-0762001. (MERI) MERCK & CO INC.
XX	PA	Davison AJ, Ellis RW, Keller PM, Lowe RS;
XX	PI	WPI; 1992-142773/18.
XX	DR	N-PSDB; AAQ24233.
XX	PT	Vaccine against varicella-zoster virus (VZV) - produced from gene encoding outer surface viral protein
XX	PS	Claim 1; Page 9; 18pp; English.
XX	CC	Chickenpox is caused by varicella-zoster virus (VZV), a member of the herpesvirus gp. VZV has five major glycoproteins on its surface which are the prods. of three different genes, gA, gB and gC. Monoclonal antibodies to gA and gB display complement-independent neutralisation and the monoclonal antibodies of gC display complement-dependent neutralisation. Cytoplasmic RNAs were prep. from VZV-infected MRC-5 cells. The RNAs encoded by the different VZV HindIII fragments were selected by hybridisation to cloned VZV HindIII DNA fragments (J.R.Ecker & R.W. Hyman, Proc. Natl. Acad. Sci. USA 79:156 (1982)) bound to nitrocellulose. These RNAs were translated in a rabbit reticulocyte lysate. The polypeptide prods. were immunoprecipitated by polyclonal monospecific guinea pig antibodies raised to gB purified by monoclonal antibody affinity chromatography. By this analysis it was found that a 100 kD in vitro translation prod. from mRNA selected by the VZV-HindIII-D fragment could be immunoprecipitated by the anti-gB antibodies which neutralise viral infectivity. The polypeptide may be used to react with human convalescent zoster sera and with monospecific antisera which neutralise viral infectivity. The polypeptide or fragments of it are useful for the prepn. of vaccines or diagnosis of VZV infection.
XX	CC	(Updated on 25-MAR-2003 to correct PF field.)
XX	SQ	Sequence 868 AA;
		Query Match 49.2%; Score 1726; DB 13; Length 868;
		Best Local Similarity 50.5%; Pred. No. 5.4e-142;
		Matches 325; Conservative 116; Mismatches 182; Indels 20; Gaps 6;
QY	31	REYSSVSVQLSEESTFLCPPPVGSGVIRLEPPKCEPRKATGWGSIATLKFENISPY 90 : : : : : : : : : : : :
Db	41	RAEIHKSQADEKPKTFYVCPPPGTGSIVRLPRTCPDYHLGNFTBGIAVVKENIAAY 100 : : : : : : : : : : :
QY	91	KFKVLTYLYNNIIQTWTGTYYRQITNRYTRDPFVSIEETDLIDGKGRCSSKARYLRNN 150 : : : : : : : : : : :
Db	101	KFKATVYVKDVIVSTAWAGSSYTQITNRVADRPVPWPVEITDTIDKFGKCSSRATYVRN 160 : : : : : : : : : : :
QY	151	VYVEAFDRDAGEQVLLKPSKFNTPSPRAHHHTTNYTVWGSGPWIRYTGTSVNCIVEMD 210 : : : : : : : : : : :
Db	161	HKVEAFEDKNPODMPLIASKNSVGSKAWHHTNDTYMWAGTGTGYTGTGSVNCIIEEVE 220 : : : : : : : : : : :

QY	211	ARSVFPYSYFAMANGDIANI	SPFYGLSPPEAAEPMPQDNFKOLDYSFMDLDRKKA	270
Db	221	ARSIFFYDFGLSTGDII	YMSPFEGCLR-DGAYREHSNYAMDRFHQFEGYQRDLTRALL	279
QY	271	SLPKRNLPLTISHFTVGDW	APKTRTRVCSMTKKWEVTEMLRATVNGRYFMARELSATFI	330
Db	280	E-PAARNEFLVTPHLTV	GNWKPKRTEVCSLVKRWREVEDVVRDEYAHNFRFTMKTLSFTFI	338
QY	331	SNTEFDNRIILGOCIKRE	AAEALEQIFRUKYNDSHVKVGHVOYFLALGGFIVAYQPVL	390
Db	339	SETNEFLNQIHLSCVKE	ERARIIINRIYTRINSHVTRGDIQTYLARGGFEVVVFQPLL	398
QY	391	SKSLAHMYRLMRDRNTD	EMLDLVNNKHAIYKKNATSLSRLLRRDIRNAPNRKIITLDDTT	450
Db	399	SNSLARLYLQELVRE-	-----NTNHS-PQKHPTNTRSRRSV-----PVELRANR	441
QY	451	AIKSTSSVOFAMQLYDHI	QTHINDMFSRTATANCELQNRELVLWHEGKINPSATSA	510
Db	442	TITTTSSVEFAMLOFTYD	HIQHEVHEMELARISSWCQLQNRERALWGLFPIINPSALAST	501
QY	511	TGLRRVAAKMGIDGVA	AVSSCTAIDAES-VTLQNSMRVITSTNTCVSRPLVLFSGYENOGN	569
Db	502	ILDQVRARIIGDVISN	CPELGSDDRILIIQNSMRVSGSTRCYSRPLISIVLSNGSGT	561
QY	570	IQGOLGENNELPLTLEA	VEPCSANHRRYFLFGSGYALFENFVKMVDAAADIQIASTFVE	629
Db	562	VEGQLGTDNELIMSRD	LLEPCVANKHRYFLFGHHVYVEDYRYREIAVHDVGMISTYVD	621
QY	630	LNLTLLDEDRETLPLSV	TKEELRDVGVLDYAEVARNOLHELK	672
Db	622	LNLTLLKDRFMPLOVY	TRDELRTDGLLDYSEIQRRNOMHSIR	664
RESULT	5			
AAW68404				
ID	AAW68404	standard; Protein; 913 AA.		
AC	AAW68404;			
XX				
DT	28-OCT-1998	(first entry)		
XX				
DE	Aujesky's disease virus glycoprotein gp.			
XX				
KW	Multivalent vaccine; pig; pathogen; respiratory disease; SIV; PRRSV; HCV;			
KW	digestive disease; Aujesky's disease virus; pseudorabies virus; vaccine;			
KW	swid herpesvirus 1; swine influenza virus; hog cholera virus; vector;			
KW	porcine respiratory and reproductive syndrome virus; glycoprotein; SIRS;			
KW	swine infertility and respiratory syndrome virus; glycoprotein; SIRS;			
KW	Actinobacillus pleuropneumoniae.			
XX				
OS	Pseudorabies virus.			
XX				
PN	FR2751224-A1.			
XX				
PD	23-JAN-1998.			
XX				
PF	19-JUL-1996; 96FR-0009338.			
XX				
PR	19-JUL-1996; 96FR-0009338.			
XX				
PA	(INMR) RHONE MERIEUX SA.			
XX				
DR	WPI; 1998-112824/11.			
DR	N-PSDB; AAV49291.			
XX				
PT	Multivalent polynucleotide vaccines against porcine pathogens			
PT	consist of at least 3 plasmids able to express protective antigens			
PT	from specified viruses			
XX				
PS	Example 8; Fig 2; 63pp; French.			
XX				
CC	The invention relates to a multivalent vaccine for protection			

CC against several pathogens, especially pathogens associated with
 CC respiratory and digestive diseases. The pathogens are especially
 CC selected from Aujeszky's disease virus, swine influenza virus (SIV),
 CC porcine respiratory and reproductive syndrome virus (PRRSV), hog
 CC cholera virus (HCV) and Actinobacillus pleuropneumoniae. The vaccines
 CC are preferably composed of polynucleotide sequences encoding 3 antigens,
 CC all as part of vectors. This sequence represents the Aujeszky's disease
 CC virus strain NIA3 glycoprotein gB. The coding sequence was subcloned
 CC into the plasmid pVR1012 to generate plasmid pAB090 for use in the
 CC vaccine.

XX
 SQ Sequence 913 AA;

Query Match 49.0%; Score 1719.5; DB 19; Length 913;
 Best Local Similarity 50.9%; Pred. NO. 2.2e-141;
 Matches 341; Conservative 105; Mismatches 183; Indels 41; Gaps 12;

QY 18 GTNSSPSTQ--NVTSSREVSVVQLSEESTFVLCPPPVGVSTVIRLEPPRKCPPEPRKATEW 75
 Db 95 GPSEAPDGEYGLDARTAVRAA--ATERDRFVCPVPPSGSVTVRLEPEQACPEYSQGRNF 152
 QY 76 GEGTALLFKNISPKFKVLYKNIQTWTGTYRQITNRYTDRTPVSEETDLID 135
 Db 153 TEGTALLFKNIAPIKFKAHYIKNVIVTTVWSGSTYAAITNRTDRVPVQVEITDVID 212
 QY 136 GKGRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPESRAWHHTTNEYTVVSGSPI 195
 Db 213 RRGKCVSAEYVRNNHKVTAEDRDENPVEVDLPSRLNALGTRAWHTTNDTYTKGAAGF 272
 QY 196 YRTGTSVNCIVEEMDARSVPYSVFAMANGDIANTSPFYGLSPPEAAEPMGYPQDNFKQ 255
 Db 273 YOTGTSVNCIVEEVARSVYDPSFALSTGLVYNSPFYGLR-EGAHGEQIGIYAPGRFQQ 331
 QY 256 LDSFESMDLRRKASLPVKRNFLLTSHFTYGMWAPKTTTRVCSMTKKVETMLR-ATV 314
 Db 332 VEHYPIIDLSRLRASESVTRNFLTPTFTVAMDWAPKTRVCSLAKWREAEEMTRDETR 391
 QY 315 NGRYFMARELSAFISNTTEFDPNRIILGOCIKREAAEQIFRTKYNDSHVKVG-HV 373
 Db 392 DGSFRFTSRALGASVSDVQLDQRVHLGDCVLRSEASEAIDAIYRRYRINSTHVLGDRP 451
 QY 374 QYFLALGFIVAYQVLSKSLAHMYLRELMDRNTDEMDLVNNKHAIYKKKNTLSRLR 433
 Db 452 EYVLARGGFVAFRPLISNELAQLYAREL-----ERLGLAG---VVGPAAPAAARRAR 501
 QY 434 RD-----INAPNRKITLDDTTAIAKSTSSVQFAMQLQFLYDHIQTHINDMFSRIA 482
 Db 502 RSPGPAGTPEPPAVNGTCHLRIT-----TGSAEARLQFTVDHIQAHVNDMLGRIA 552
 QY 483 TAWCELQRELVLWHEGKINPSATASATILGRRVAAKMLGDVAVSSCTAIDAESVTIQN 542
 Db 553 AAWCELQKDRTLWSEMRLNPSAVATAALGORVCARMLGDVMAISRCVEVRG-GVYYQN 611
 QY 543 SMRVITSTNTCYSRPLVFSYGENOGNIGOLGENNELLPTLEAVEPCSAHNRFFYFGS 602
 Db 612 SMRVPGERTCYSRPLVTFEH-NGTGVIEGGDGDNDNELLISRDLEPCTGNHRRYFKLGS 670
 QY 603 GYALFENTYFVMDAADIQIASTFVELNLTLLEDEITPLSVYTKBELRDVGVLDYAEV 662
 Db 671 GYVYEDYNYVRMVEVP--ETISTRTVNLNLLEDEITPLSVYTKBELRDVGVLDYAEV 728
 QY 663 ARNQLHELK 672
 Db 729 QRNQLHALK 738

RESULT 6

AAP50035

ID AAP50035: standard; Protein; 845 AA.

XX

AC

XX

XX

DT 25-MAR-2003 (updated)

DT 07-SEP-1991 (first entry)
 DE N-terminal sequence of the pseudorabies virus (PRV) gII protein.
 KW Sub-unit vaccine; antigen; immunoreactive determinant.
 XX Pseudorabies virus (PRV).
 FH Key Location/Qualifiers
 FT Modified-site 151..153 /label= potential glycosylation site
 FT Modified-site 261..263 /label= potential glycosylation site
 FT Modified-site 441..443 /label= potential glycosylation site
 FT Modified-site 516..518 /label= potential glycosylation site
 FT Modified-site 573..575 /label= potential glycosylation site
 FT Modified-site 633..635 /label= potential glycosylation site
 FT Modified-site 697..699 /label= potential glycosylation site
 XX EP162738-A.
 XX 27-NOV-1985.
 XX 09-APR-1985; 85EP-0400704.
 XX 03-APR-1985; 85US-0719773.
 XX 09-APR-1984; 84US-0598073.
 XX (MOLE-) MOLECULAR GENETICS INC.
 PA Robbins AK, Watson RJ, Enquist LW;
 XX WPI; 1985-298083/48.
 DR N-PSDB; NAM50036.
 XX Prodn. of pseudorabies virus sub-unit vaccines - useful for
 PT conferring protection against the virus infections and for
 PT diagnosis esp. with sheep sera
 XX Example; Fig 13; 153pp; English.
 PS The inventors claim a recombinant vector comprising a DNA sequence
 CC coding for an immunoreactive and antigenic determinant of a
 CC pseudorabies virus protein, and a polypeptide having an
 CC immunoreactive and antigenic determinant of a pseudorabies virus
 CC protein. Vaccines contg. the polypeptide are obtd. economically and
 CC in large amts., for use in conferring protection against
 CC pseudorabies virus.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 845 AA;

Query Match 48.9%; Score 1715.5; DB 6; Length 845;
 Best Local Similarity 50.7%; Pred. No. 4.3e-141;
 Matches 340; Conservative 105; Mismatches 184; Indels 41; Gaps 12;

QY 18 GTNSSPSTQ--NVTSSREVSVVQLSEESTFVLCPPPVGVSTVIRLEPPRKCPPEPRKATEW 75
 Db 95 GPSEAPDGEYGLDARTAVRAA--ATERDRFVCPVPPSGSVTVRLEPEQACPEYSQGRNF 152
 QY 76 GEGTALLFKNISPKFKVLYKNIQTWTGTYRQITNRYTDRTPVSEETDLID 135
 Db 153 TEGTALLFKNIAPIKFKAHYIKNVIVTTVWSGSTYAAITNRTDRVPVQVEITDVID 212
 QY 136 GKGRCSSKARYLRNNVVEAFDRDAGEKQVLLKPSKFNTPESRAWHHTTNEYTVVSGSPI 195
 Db 213 RRGKCVSAEYVRNNHKVTAEDRDENPVEVDLPSRLNALGTRAWHTTNDTYTKGAAGF 272

196 YRTGTSVNCIVEEMDARSVPYSYFAMANGDIANISPFYGLSPPEAAEPMPGPDNFQ 255
 273 YHTGTSVNCIVEEARSVPYDPSFALSGDIVMSPFFGLR-EGAHGEHIGVAPGRFOQ 331
 256 LDSYFSDLDKRRKASLPVKRNLITSHFTVGDWAPKTRVCSMTKWEVTEMLR-ATV 314
 332 VEHYPIDLSDLRASESVTRNFRPHFTVADWAPKTRRVCSLAKWEAEWTRDER 391
 315 NGRYRMARELSATFISNTTEPDNRIILGQCIREAEAAIQIFRTKYNDSHVKG-HV 373
 392 DGSFRTSALGASFVSDVTQDLQKRVHLGDCVLRASEAIDAIYRRYNSHVLGDRP 451
 374 QYFLAGGFIVAYQPVLSKSLAHMYLRLMRDNRDTEMLDLVNNKHAIYKKNATSLRLR 433
 452 EYVLARGGFVAFRPLISNELAQLYREL-----ERGLAG----VVGPAAPAAARAR 501
 434 RD-----IRNPNRKITLDDTTAISTSSVQFAMQLQFLYDHIQTHINDMFSRIA 482
 502 RSPGPAGTPEPPAVNGTGHRLIT-----TGSAEFARLQFTYDHIQAHVNDMLGRIA 552
 483 TAWCELQNLRELWHEGKINPSATASATLGRVAAKMLGDVAVSSCTADAESVTIQN 542
 553 AAWCELQNKDRFLWSEMSRLNPSAVATAALGORVSARMIGDVAISRCVEVRG-GVYVQN 611
 543 SMRVITSTWTCYRPLVFLPSYGENQNGIQQGNNELLPTLEAVEPCSANHRRYFLGS 602
 612 SMRVPGERTCYRPLVTFEH-NGTGVTEGQDGDNNELLISRDLEPCTGNHRRYFLGS 670
 603 GYALFENTNFVKMDAADIQASTFVELNLTLEDREILPLSVYTKBELROGVLDVAEV 662
 671 GYVYEDYNYRMVEVP--ETISTRVLNLTLEDREPLFLEVYTRTELADTGLDYSEI 728
 663 ARNQLHEK 672
 729 QRNQLHALK 738

RESULT 7
 AAR92746
 ID AAR92746 standard; protein; 891 AA.
 XX
 AC AAR92746;
 XX
 DT 11-MAY-1996 (first entry)
 XX
 DE B virus gB glycoprotein.
 XX
 KW Herpes simian monkey B virus gB glycoprotein; UL27; immunoassay;
 KW diagnosis; herpes B virus.
 XX
 OS Herpes simian monkey B virus.
 XX
 PN US5487969-A.
 XX
 PD 30-JAN-1996.
 XX
 PF 01-APR-1993; 93US-0042747.
 XX
 PR 01-APR-1993; 93US-0042747.
 XX
 PA (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.
 XX
 PI Black D, Eberle R, Hilliard J, Scinicariello F;
 XX
 DR WPI; 1996-105220/11.
 DR N-PSDB; AAT16474.
 XX

Detection of herpes B virus by PCR amplification of sample DNA - to
 detect a specific herpes simian monkey B virus DNA segment.
 Claim 1; Column 19-24; 22pp; English.

CC The herpes simian monkey B virus proteins, such as the gB
 CC glycoprotein (UL27), have immense potential use in the development
 CC of serological immunoassays, which can specifically detect virus
 CC antigens and/or antibodies to B virus. One approach is to
 CC synthesize peptides which, based on the properties of the predicted
 CC protein sequence, are likely to be immunologically active. Such
 CC peptides can be used as substrate antigens in immunoassays to detect
 CC serum antibodies which recognize this specific peptide sequence.
 CC Synthetic peptides may also be used to produce antibodies against
 CC specific regions of the gB glycoprotein which are unique to one
 CC virus. These can then be used to develop virus-specific
 CC immunoassays for differentiation of B virus from other primate
 CC alpha-herpes virus and for identification of antibodies directed
 CC against B virus in primate serum samples.
 CC
 XX Sequence 891 AA;

Query Match 48.4%; Score 1699; DB 17; Length 891;
 Best Local Similarity 49.2%; Pred. No. 1.3e-139;
 Matches 322; Conservative 99; Mismatches 210; Indels 24; Gaps 2;
 QY 18 GTNSPSTQNVTSREVSVSSVLSSEESTFYLCPPPGVSTVIRLEPPKCPKATEWGE 77
 DB 72 GTNASVEAGHATLRENDIKALGDATFYVCPPTGATVYQFEPQPCPRAPHGQNYE 131
 QY 78 GIALFKENTISPYKFKVLYKNIQTWTGTTTQITNRYTDTPTVSTEEITDLIDGK 137
 DB 132 GIATVIFENIAPYKFKATMYKDVTVSQVWFGHYSQFMGIFEDRAPVPEEVIDKINAR 191
 QY 138 GRCSKARYLNNVVVEAFDORAGEKQVLLKPSKFNTPESRAHHTTNETYVMGSPWYR 197
 DB 192 GVCSTAKYVNNMESTAFHRDDSDMKLKPAAKATRTSRGWHITTLKYNPSIEAFHR 251
 QY 198 TGTSVNCIVEEMDARSVPYSYFAMANGDIANTSPFYGLSPPEAAEPMPGPDNFQ 257
 DB 252 YGTVNCIVEEARSVPYDEFVLAGDFVYSPFYGYR-DGAHAHTAIAADRFQVD 310
 QY 258 SYFSMDLDKRRKASLPVKRNLITSHFTVGDWAPKTRVCSMTKWEVTEMLRATVNGR 317
 DB 311 GYTERDLSTGRRASTPATRNLLTPKFTVGDWAPKRPVSCTLTKEQVDEMLRAEYGPS 370
 QY 318 YFMARELSATFISNTTEPDNRIILGQCIREAEAAIQIFRTKYNDSHVKGHVOYFL 377
 DB 371 FRESSSALSTFTTNETEYALSVDGDCVGREAREAVDRIFLURYNGTHVKGVOYFL 430
 QY 378 ALGGFIVAYQPVLSKSLAHMYLRLMRDNRDTEMLDLVNNKHAIYKKNATSLRLRDIR 437
 DB 431 ATAGFLIAYQPLLSNGLVELYVRELLREQEGRP-----GDAA 467
 QY 438 NAPNRKITLDDTTAISTSSVQFAMQLQFLYDHIQTHINDMFSRIAATWCELQNLRELW 497
 DB 468 ATPKPSADPPDVERIKTSSVEFARLQFTYDHIQAHVNDMLGRIAATWCELQNLRELW 527
 QY 498 EGKINPSATASATLGRVAAKMLGDVAVSSCTADAESVTLONSMRVITSTNTCYSRP 557
 DB 528 EARKLNPAIASATVGRVRSARMLGDVMAVSTCVPTPDVIMQNSMRVPARPGTCYSRP 587
 QY 558 LVLFYSGENQNGIQQGNNELLPTLEAVEPCSANHRRYFLFGSGYALFENYFVKWD 617
 DB 588 LVSFYEGGPLVEGQLGEDNEIRLERDALEPCTVGHRRYFTFGAGYVYFEDYAYSHQIG 647
 QY 618 AADIQIASTFVELNLTLEDREILPLSVYTKBELROGVLDVAEVARRNQLHEK 672
 DB 648 RADVTVTSTFINLNTLEDREHFEVPLEVYTRQEKDGLDYTEVQRRNQLHALR 702

RESULT 8
 AAW70293
 ID AAW70293 standard; protein; 891 AA.
 XX
 AC AAW70293;
 XX
 DT 06-NOV-1998 (first entry)

Qy	18	GTNSPSTONVTSREVVSVOQLSEESTFYLCPPPGSVTVIRLEPPKPCPEPRKATWGE	77
Db	97	GDDAASPDNSTDVRAALRLAQAGENSREFFVCPPPSGATVVR LAPARPCEYGLGRNTE	156
Qy	78	GIAILFENISPKFKVTLIYKNIIOTTTTGTVTRQITNRYTORTPVSEIEETDLIDGK	137
Db	157	GIGVYIKENIAPYTKAYIYKKNVITVTTWAGSTYAAITNQYTORVPVGMGEITDLVDKK	216
Qy	138	GRCSSKARVLNNVYVEAFDRDAGEKOVLLKSPKSFENTPESRAWHTTNETTVVGSWPIYR	197
Db	217	WRCLSKAEYLRGSRKVAFDRDDDPWEAPLAPRLASAPGVGWHTTDDVVTALGSLAGLR	276
Qy	198	TGTSVNCIIVEEMDARSVPYPSYFAMANGDIANISPFYGLSPPEAAABPMGYPODNFKQLD	257
Db	277	TGTSVNCIIVEEYARSVPYDYSFALSTGDIILYSPFYGLR-EGAHREHTSYSPEREQIE	335
Qy	258	SYFSMDLKKRRKASLPVKRNFLITSHFTVGNWDAPKTTRVCSMTKKWEVTEMLRATVNGR	317
Db	336	GIYKRDMA7GRRLKEPVSRNFLTQHTVTVANDWVPKRNKVCSLAKHREADEMLRDSESRN	395
Qy	318	YRFMARELSATPISNTBFDPNRIILGOCIKREAAAEIQIFRTKYNDSHVKVGHVQYFL	377
Db	396	FRFTARSLSATFVSDSHTFALQNVPLSDCVTEEEAAAVERYERYNGTHVLSGSLETYL	455
Qy	378	ALGCFVAYOPVLSLSLAHMYLRELMRNRNDEMULDVNKNHAIYKKNATLSRLRDRTR	437
Db	456	ARGGFVAFRPMLSNELAKYLQELARNSGTLEGFAA-----AAPKPGP---RRARARAP	508
Qy	438	NAPNRRKITL-----DDTTAIKTSYSQVAFMLQFLYDHIQTHINDMFSRATAMCELQN	490
Db	509	SAPGGPGAANGPAGDGDAGGRVTVSSAEFAALQFTYDHIQDHVNTFMSRLATSWCLLQN	568
Qy	491	RELVLWHEGIKINPSATASATLGRRVAAKMLGDVAVSSCTAIDAGSVTLQNSMRVITST	550
Db	569	KERALWAEAAKLNPSSAAASALDRRAARMLGMDAMAVTYCHELGEGRVFIENSMR--APG	626
Qy	551	NTCYSRPLVLSYSGNQGNIQGLGENNELPTLEAVEPCSMNRHYFFLFGSGYALFENY	610
Db	627	GVCYSRPPVFAFGNESEPVESQGLGEDNELLPGLRELVEPCTANHKKYFRFGADYVYENY	686
Qy	611	NEVKMVDAAADTQIASTFVELNLTLLEDBREILPLSVYTKKEELROGVGLDYAEVARRNQLHE	670
Db	687	AVRRVPLAELEVISITFVDNLNLTLQDREFLPLEVYTRAEIADTGLDYSIEQRRNQLHE	746
Qy	671	LK	672
Db	747	LR	748

RESULT	10
AAR411343	
ID	AAR411343 standard; Protein; 933 AA.
XX	
AC	AAR41343;
XX	
DT	18-FEB-1994 (first entry)
XX	
DE	Bovine herpesvirus type I gI glycoprotein.
XX	
KW	BHV-1; Bovine herpesvirus; vaccine; neutralising epitope;
XX	glycoprotein; coat protein; gI; herpes; virus.
XX	
OS	Bovine herpesvirus.
XX	
FH	Key
FF	Location/Qualifiers
Cleavage-site	505
FT	
FT	/note= gIa precursor protein cleaved to give gIb
FT	and gIc
FT	
Region	767..828
FT	
FT	/label= Putative transmembrane region.
XX	
PN	CA2057387-A.

```

XX      12-JUN-1993.
PD
XX
XX      11-DEC-1991;    9ICA-2057387.
PF
XX
XX      11-DEC-1991;    9ICA-2057387.
PR
XX
XX      (VETE-) VETERINARY INFECTIOUS DISEASE.
PA
XX
XX      Babluk L, Fitzpatrick D, Van DER HURK S, Zamb T;
PI
XX
XX      WPI: I993-273267/35.
DR
XX      P-FSDB: AAR41343.
DR
XX
XX      Recombinant bovine herpes type 1 proteins gI, gIII and gIV - and
PT      DNA encoding them, for use in vaccines against e.g. shipping
PT      fever
PT
XX
XX      Claim 2; Figure 5; 154pp; English.
PS
XX
XX      The bovine herpesvirus gI glycoprotein (BHV-1) is used in the
CC      construction of vectors which are subsequently used to transform
CC      host cells. gI produced by these cells can then be used to produce
CC      subunit vaccines comprising one or more neutralising epitopes of the
CC      gI glycoprotein. The subunit vaccines are used to protect cattle
CC      from disease, especially shipping fever which is a complex syndrome
CC      which often includes infection by BHV-1. They are substantially
CC      more protective than previous killed or live attenuated virus
CC      vaccines.
XX
XX      Sequence       933 AA;
SQ
Query Match          48.3%; Score 1695.5; DB 14; Length 933;
Best Local Similarity 49.8%; Pred. No. 2.9e-139;
Matches 330; Conservative 102; Mismatches 213; Indels 17; Gaps 5;

Qy      18 GTNSPSTQNTVREVYSSVLSEESTFFILCPPPGVSTVIRLEPPPKCDEPRKATWGE 77
Db      97 GDNAASPDNDTVRAURLQAAGENSRRFFVCPPPSGATVVVRLLAPARCPCPEYGLGRNTE 150

```

Query Match	48.38;	Score	1695.5;	DB	14;	Length	933;
Best Local Similarity	49.8%;	Pred.	No. 2.9e-139;				
Matches	330;	Conservative	102;	Mismatches	213;	Indels	17;
Gaps	5;						
QY	18	GTNSSPQTQNTSVREVSSVQLSEESTFYFLCPPPGVSTVIRLEPPKRCDEPRKATWGE	77		:	:	:
Db	97	GDDAASPDNDTVRAURLAQAGENSRRFFVCCPPSGATVVRAPARPCPEYGLGRNTE	156				:
QY	78	GIAIFKENISPYKFKVLYLYKNIQTTTGTGTYRQITNRYITDRTVSTEEITDLIDGK	137				
Db	157	GIGVIYKENIAPYTFKAIYIYKKNIVITTTWAGSTYAAITNQYTDTRVPVGMGEITDIDYKK	216				
QY	138	GRCCSKARYLRNNVYEAPORDAGEKOVLLKPSKFNTPESRAHHTTNETTVTVMGSPWIYR	197				
Db	217	WRCLSKAEYLRSGKVVAFDRDDDPWPAKLPAKLRSAPGVRGVHTTTDDVYTLGSAQLYR	276				
QY	198	TGTSVNCIVEEMDARSVPFYSYFAMNGDITANISPFYGLSPPEAAAEPMGYPODNFKOLD	257				
Db	277	TGTSVNCIVEEARSVYPYDSFALSIGDIIYMSFPYGLR-EGAHREHTTSYSPERPOQIE	335				
QY	258	SYFSMDLDRKKASLPVKRNFILTSHTPTQWMDNAPKTRVCSMTKKWEKTEMLRATVNGR	317				
Db	336	GYYKRDMAIGRRLEKPEVSRNFLRQHTVAVMDVWPKKKNVCSLAKWREADMLRDESGRN	395				
QY	318	YRWMARELSATFISNTEFPDPNRIILGCQIKREAEALQEIFTKYNDSHVKVGHVOYFL	377				
Db	396	FRTARSLSATFVSDSHTFALQNVPLSDCVIEEAAEAVERYRNGHTVHSGSLFETYL	455				
QY	378	ALGGFIVAYOPVLKSLAHMLRELNRDNTDEMLDLVNNKHAIYKKNATSLSLRRDRIR	437				
Db	456	ARGGFVVAFRPMLSNELAKIYLQELARSNCTLSGLFAA----AAPRGP---RRARRAAP	508				
QY	438	NAPNRKLTITL-----DDTTAKTSSTGVFAMLOFTYDHIQTHINDMFSKIATAWCELQN	490				
Db	509	SAPGPGGAANGPAGDGDAGGRVTVTSSAEFAALQFTYDHIQDVNTVMFSRLATSWCLLQN	568				
QY	491	RELVLWHEGIKINPSATLGRYRAAKMLGDVAAVSCTAIDAESVTLONSMRVITST	550				
Db	569	KERALWAEAAKLNPSAAASALDRRAARMLGDMAAVTYCHEUGEGRVFIENSMR--APG	636				
QY	551	NTCYSRPLVLFVSYGNOGNIQOLGENNELLTLEAVEPCSAHHRIFYFLGSGYALFENY	610				


```

Db 441 GGFLIAYQPLLSNTLAEIYVREYREQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNRKTLTDDTTAISTSSVQFAMQLQFLYDHIOTINDMFSTRATAWCELQNLRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHIOHRYNDMLGRIAVAWCELQNHLETLWNEA 539
QY 500 IKINPSATASATLGRVAAKMLGDVAAYSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAISATVGRVRSARMLGDVMAVSTCVVPADNVIVQNSMRVSSRPGTCYSRPLV 599
QY 560 LFSYGENQGN-IQGLQGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFKWYDA 618
Db 600 SFRY-EDOGPLIEGOLGNNELRLTRDALEPCTVGHRRYFIFGGYVYFEYAYSHQLSR 658
QY 619 ADIQIASTFVELNLTLLEDREILPLSVYTKBELRDVGVLDAEAVARRNQLHELK 672
Db 659 ADVTTVSTFIDLNTMLEDHFPVLEVYTRHEIKDGLDVTQVRRNQLHDLR 712

RESULT 14
AAW00376
ID AAW00376 standard; Protein; 904 AA.
XX AC AAW00376;
XX DT 24-JAN-1997 (first entry)
XX DE HSV-2 glycoprotein B.
XX KW HSV-2; glycoprotein B; vaccine; diagnosis; immunoassay.
XX OS Herpes simplex virus type 2.
XX FH Key Location/Qualifiers
FT Peptide 1..22
FT Peptide /label= Sig_peptide
FT Peptide 18..75
FT Domain /label= Dominant_type_specific_epitope
FT FT 23..744
FT FT /label= Extracellular_domain
FT FT 745..798
FT FT /label= Transmembrane_domain
FT FT 799..804
FT FT /label= Cytoplasmic_domain
FT FT 819..904
FT FT /label= Cross-reactive_antigenic_epitope
XX WO9632962-A1.
XX PD 24-OCT-1996.
XX PF 19-APR-1996; 96WO-US05316.
XX PR 21-APR-1995; 95US-0426604.
XX PA (UYNE-) UNIV NEW MEXICO STATE.
XX PI Bell R, Goade DE, Jenison S;
XX WPI; 1996-485557/48.
XX DR
XX PT New type-specific and cross-reactive Herpes Simplex Virus epitope(s)
XX PT - used for diagnosing HSV-1 and/or HSV-2 infections and in vaccines
XX PT for producing neutralising antibodies
XX PS Disclosure; Fig 5; 37pp; English.
XX CC Herpes simplex virus type 2 (HSV-2) glycoprotein B (gB2) (AAW00376)
XX CC and HSV-1 gB1 (AAW00375) include type-specific and cross-reactive
XX CC antigenic epitopes. Recombinant type-specific epitopes can be used
XX CC in immunoassays to distinguish between HSV-1 and HSV-2 infections.
XX CC Recombinant cross-reactive epitopes are useful for the diagnosis of

```

```

CC both HSV-1 and HSV-2 infections. Both types of epitope can be used
XX in vaccines.
SQ Sequence 904 AA;
Query Match 48.2%; Score 1692; DB 17; Length 904;
Best Local Similarity 49.1%; Pred. No. 5.6e-139;
Matches 321; Conservative 115; Mismatches 194; Indels 24; Gaps 7;
QY 20 NSSPSTQNTREVSVSSVQLSEESTFYICPPVPGSTVIRLPPRKCPEPRKATWEGEI 79
Db 82 NATVAAGHATLAHLREIKVENADAQFYVCPPTGATVVQFPQRCRPRPGQNYTEGI 141
QY 80 ALLFRENISPKFKVLYLKNIIQITWTGTITNRYDTRTPSVSEETEDLDIDGKR 139
Db 142 AVFKEKNIAPYKFKATMYKQVTVSQVWFHGRYSQFMGIFEDRAPVPEEVIDKINAKVG 201
QY 140 CSSKARYRNNVYVEAFDRDAGEKOVLLKPKSFNTPEPRAWHTTNETYVWGSPIYRTG 199
Db 202 CRSTAKYVRNNMETTAFHRDOHETDMLKPAKAVATRTSGWHTTDLKYNPSRVEAFHFG 261
QY 200 TSVNCIVEEMDARSVPYYSYFAMANGDIANISPPYGLSPPEAAABPMGYPDQNFQLODSY 259
Db 262 TVNCIVEEVDARSVPYDEFLATGDFVYNSPFYGYR-EGSHTETSVAADRFFQVDGF 320
QY 260 FMDLDKRRKASLPVKRNFELITSHFTVGWDWNAKPTTRVCSMTKWKVEVTEMLRATVNGRYR 319
Db 321 YARDLTTKARATSPTRNLLTTPKFTVAMDWYKRPVACTMTKWOEVDMLRAEYGGSPR 380
QY 320 FMARELSATFISNTTEFDPNRIILGQCIKREAAAEIQIFRTKYNDSHVKVGHVOYFLAL 379
Db 381 FSSDAISITFTTNLTQYLSRVDLGCIGRDAREALDRMFARKYNATHIKVGPQPYLAT 440
QY 380 GGFIVAYQPLVSKSLAHMYLRELMRDNRDTEMLDLVNNKHAIYKKNATSLSLRLRDIRNA 439
Db 441 GGFLIAYQPLLSNTLAEIYVREYREQ-----DRK-----PRNATP-----APLREA 482
QY 440 PNRKTLTDDTTAISTSSVQFAMQLQFLYDHIOTINDMFSTRATAWCELQNLRELVLWHEG 499
Db 483 PSANASVE---RIKTTSSIEFARLQFTYNHIOHRYNDMLGRIAVAWCELQNHLETLWNEA 539
QY 500 IKINPSATASATLGRVAAKMLGDVAAYSSCTAIDAESVTLQNSMRVITSTNTCYSRPLV 559
Db 540 RKLNPNAISATVGRVRSARMLGDVMAVSTCVVPADNVIVQNSMRVSSRPGTCYSRPLV 599
QY 560 LFSYGENQGN-IQGLQGENNELPTLEAVEPCSANHRRYFLFGSGYALFENYFKWYDA 618
Db 600 SFRY-EDOGPLIEGOLGNNELRLTRDALEPCTVGHRRYFIFGGYVYFEYAYSHQLSR 658
QY 619 ADIQIASTFVELNLTLLEDREILPLSVYTKBELRDVGVLDAEAVARRNQLHELK 672
Db 659 ADVTTVSTFIDLNTMLEDHFPVLEVYTRHEIKDGLDVTQVRRNQLHDLR 712

RESULT 15
AAW34552
ID AAW34552 standard; Protein; 904 AA.
XX AC AAW34552;
XX DT 26-FEB-1998 (first entry)
XX DE Herpes simplex virus type 1 glycoprotein B.
XX KW Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
XX KW membrane protein; virus-specific glycoprotein;
XX KW transmembrane anchor region.
XX OS Herpes simplex virus type 1.
XX FH Key Location/Qualifiers
FT Domain 726..795

```

